

### De toetssteen des aanstoots

Een proef op de som van Van Houwelingen,  
Tazelaar en Verbeek

Jan Reijnders

#### 1. Inleiding

In het themanummer van de Sociologische Gids over Werkloosheid (XXXI, 1984, p. 6-23) gaan Van Houwelingen, Tazelaar en Verbeek in op een mogelijk verband tussen de economische situatie, voor zover geïndiceerd door het werkloosheidspercentage, en gezondheid respectievelijk sterfte. In dit verband richten zij hun boze oog op het werk van de Amerikaan Harvey M. Brenner, voor wie het bestaan van een dergelijk verband nauwelijks meer aan twijfel onderhevig lijkt te zijn (Brenner, 1971, 1975, 1977, 1979, 1983).

Brenners' werk wordt bepaald niet gekarakteriseerd door helderheid, eenduidigheid en systematiek. Voor een dergelijk gebrek zijn in principe verschillende verklaringen te geven. Gezien zijn stijl van werken ligt het voor de hand te veronderstellen dat hij gedreven door zijn streven om empirische ondersteuning voor zijn hypothese te vinden, de ambiguïteit enigszins als instrument gebruikt. Klaarblijkelijk gaan ook Van Houwelingen c.s. van een dergelijke verklaring uit. Vanuit dit standpunt bezien is het grootste deel van hun kritiek dan ook vrij voor de hand liggend.

Wat hun eigen bijdrage betreft legitimeren de auteurs hun interventie op twee gronden:

1. zij bekritisieren Brenners' toetsingsmethode en stellen hier een andere tegenover (p. 10-11);
2. zij plaatsen vraagtekens bij de veronderstelde universaliteit van Brenners' bewering en stellen er een toets tegenover die gebaseerd is op nederlandse data (p. 9-10).

Beide gronden zijn op zich legitiem. Men kan zich echter afvragen of juist deze combinatie wel zo gelukkig gekozen is. Het is althans achteraf moeilijk om het verkregen resultaat juist te beoordelen. Is hun resultaat – namelijk dat een dergelijke samenhang niet bestaat – een attribuut van de methode, van het specifieke toepassingsgebied of van beide?

In deze reactie wil ik me beperken tot slechts één aspect van het laatste probleem, namelijk het vraagstuk van de methode. De vraag is: leidt de toets die Van Houwelingen c.s. voorstellen tot betrouwbare uitspraken over het door Brenner gepostuleerde verband of niet? Hiertoe besteed ik allereerst aandacht aan de datatransformaties die de auteurs uitvoeren. Daarbij valt te constateren dat de resultaten hoe dan ook uiterst voorzichtig geïnterpreteerd dienen te worden in verband met de dataverontreiniging die de transformaties met zich meebrengen. Vervolgens bespreek ik de methode die de auteurs toepassen en de voorwaarden waaronder hun toetsingscriterium daarmee in overeenstemming is. Ik wil daarbij aantonen dat de toetssteen, gegeven de impliciete vooronderstellingen bij de datatransformatie, niet consistent is met de eigenschappen van de methode die zij toepassen.

## 2. Datatransformaties. Het model en de impliciete vooronderstellingen

Het bevestigen of ontkennen van een verband tussen werkloosheid en sterfte is wezenlijk een vraagstuk van tijdreeksanalyse. In deze context is slechts dan van een verband sprake als een systematisch patroon dat zich in het tijdsverloop van de onafhankelijke variabele voordoet, zijn weerspiegeling vindt in het tijdsverloop van de afhankelijke variabelen. Het opnemen van de term 'systematisch patroon' refereert impliciet aan het bestaan van een tweede tijdreekscomponent: het 'niet-systematische patroon', dat ik verder zal aanduiden met de term *random* sequentie of *random* component. In de regel wordt het systematische patroon verder onderverdeeld in een *trend* en een systematische component. We komen zodoende tot de eenvoudigste vorm van een tijdreeksmodel dat er als volgt zou kunnen uitzien:  $Y_t = T_t + S_t + R_t$ . Hierbij wordt de *trend* ( $T_t$ ) opgevat als een monotone ontwikkeling in één bepaalde richting. De systematische component ( $S_t$ ) wordt verondersteld een min of meer regelmatige herhaling van een soortgelijk patroon te representeren en kan in abstracto als een periodieke functie van de tijd worden voorgesteld<sup>1</sup>. De *random* component ( $R_t$ ) tenslotte wordt opgevat als een opeenvolging van onderling onafhankelijke *random* waarden. Hoewel de auteurs zich daarover niet uitspreken, kan men uit hun operationalisatie opmaken dat zij impliciet van een soortgelijk model uitgaan (p. 12).

Op basis van een dergelijk model zou een toets voor een verondersteld verband formeel voorafgegaan moeten worden door eliminatie van zowel de *trend* als van de *random* component. Het eerste om illusoire *trend*-correlatie te vermijden, het tweede om te voorkomen dat de toetsingsresultaten onevenredig worden beïnvloed door toevallige factoren en waarnemingsfouten. Vanuit dit standpunt bezien, is de kwaliteit van een bepaalde toets afhankelijk van de mate waarin en de precisie waarmee men in staat is de ongewenste invloeden uit te schakelen, dan wel onder controle te brengen.

In de praktijk zijn beide eliminatieprocedures een delicate aangelegenheid. Wat de *trend* betreft bestaan er eigenlijk geen exacte apriori criteria op grond waarvan eliminatie zou moeten plaatsvinden. Bovendien is de ogenschijnlijke *trend* in veel gevallen kritisch afhankelijk van de lengte van het geanalyseerde tijdsinterval waardoor het hanteren van globale 'praktische' criteria verraderlijk kan zijn (Melnyk, 1969, p. 17; Reijnders, 1984, p. 415-417). Nog problematischer, zo niet onmogelijk, is het elimineren van de *random* component. Men zou *datasmoothing* kunnen overwegen (zie echter paragraaf 3) of een indirecte controle waarbij men bijvoorbeeld op grond van de 'variate difference'-methode (Tintner, 1940) een schatting van de variantie van de *random* component maakt. Dit laatste is slechts mogelijk als men aanneemt dat de systematische component (inclusief de *trend*) exact wordt beschreven door een polynoom van de graad  $n$  (Kendall en Stuart, 1968, p. 385).

Van Houwelingen c.s. nemen aan dat het probleem wordt opgelost door twee afzonderlijke datatransformaties uit te voeren:

1. de 'complexe operationalisatie' waarbij de oorspronkelijke reeksen worden omgezet in afwijkingen ten opzichte van een gewogen<sup>2</sup> 9-punts voortschrijdend jaargemiddelde (p. 12);
2. de 'simplele operationalisatie' waarbij de oorspronkelijke reeksen worden omgezet in hun eerste-differenties (*idem*).

Hoewel de schrijvers impliceren dat er een duidelijk verschil bestaat tussen beide operationalisaties, behoren beide in feite tot dezelfde klasse. Beide gaan immers van een voortschrijdend gemiddelde uit. Het onderscheid tussen de twee heeft slechts betrekking op de 'lengte' van het respectieve voortschrijdend gemiddelde (9 resp. 2 jaar) en op de samenstelling van de wegingsvector.

Een voortschrijdend gemiddelde is in principe een 'analoog'<sup>3</sup> filter dat tot doel heeft periodieke componenten te elimineren waarvan de periode korter dan wel gelijk is aan de lengte van het voortschrijdend gemiddelde. Het 9-punts filter dat Van Houwelingen c.s. hanteren elimineert dus in principe periodieke componenten waarvan de periode kleiner dan of gelijk is aan 9 jaar. Door vervolgens de afwijkingen van de oorspronkelijke reeks ten opzichte van het 9-punts gemiddelde te bepalen, definiëren zij een residu dat in principe slechts die systematische component bevat waarvan de 'periode' 9 jaar is of korter. Het eerste-differentiefilter geeft vergelijkbare resultaten. Het onderdrukt de amplitude van een systematische component waarvan de periode langer is dan 6 jaar (Yule, 1921, p. 506). De werking van het eerste-differentiefilter verschilt van de werking van het 9-punts filter in die zin dat differenties het effect van korte-termijnschommelingen accentueren.

Uit deze datatransformaties zijn voorlopig twee belangrijke conclusies te trekken:

1. De definitie van de residuen waarop de toets van het verband tussen werkloosheid en sterfte wordt gebaseerd, bevat impliciet een specifieke hypothese omtrent de aard van dit verband. Het wordt namelijk geacht beperkt te zijn tot die systematische componenten waarvan de 'periode' 9 jaar of korter is. Vergelijken we dit met de classificatie van bewegingstypen die in de economische conjunctuurleer gangbaar is (Van Duijn, 1983, p. 3-19), dan blijkt dat slechts de invloed van de korte conjunctuurcyclus (minor business cycle) en van (een deel van) de Juglar-cyclus (major business cycle) met een lengte van 4 à 5 respectievelijk 8 à 12 jaar in de beschouwing betrokken worden. De zogenaamde Kuznets- en Kondratieff-golven met een lengte van 15 à 20 jaar en 50 à 60 jaar worden klaarblijkelijk tot het domein van de *trend* gerekend. De mogelijk in deze sfeer aanwezige verbanden worden bijgevolg buiten beschouwing gelaten.

2. Aangezien elke individuele waarde van de tijdreeks een *random* fractie bevat, de *random* fracties onafhankelijk zijn en de filterprocedures slechts bewegingen elimineren die een enigszins glad verloop hebben, blijft de *random* component in principe<sup>4</sup> in de gefilterde reeksen aanwezig. Afhankelijk van het relatieve gewicht van deze component zal een toets op het residu de suggestie wekken dat geen verband aanwezig is of dat slechts van een zwak verband sprake is.

### 3. Datatransformatie en dataverontreiniging

In het voorgaande ging ik er nog van uit dat het voortschrijdend gemiddelde perfect werkte, in de zin dat zijn effect exact aan zijn doel beantwoordde en dat het geen belangrijke bijwerkingen vertoonde.

In de praktijk is het juist het complete gebrek aan perfectie dat het voortschrijdend gemiddelde een slechte reputatie bezorgde. Kendall en Stuart (1968, p. 375) maken hierover een behartigenswaardige opmerking: 'If we are concerned chiefly to describe a broad trend in the data, and are not particularly interested in short-term and residual effects, one type of moving average may be adequate. But if we want to remove the trend in order to study the residuals, such a type may be quite inappropriate; and indeed, for some purposes, we may well question whether it is safe to eliminate trends by a moving average at all' (cursivering van mij).

De reden voor deze en soortgelijke waarschuwingen is te vinden in de bijwerkingen die voortschrijdend gemiddelden vertonen. Waar het gaat om de systematische component worden de bijwerkingen bepaald door de eigenschap dat het voortschrijdend gemiddelde per definitie slechts geschikt is om bewegingen te elimineren die een constante periode en een constante amplitude hebben, voor zover de periode exact overeenkomt met de lengte van het gemiddelde of één van zijn hogere harmonischen<sup>5</sup>. In het geval waarin de systemati-

sche component bestaat uit een cyclische beweging die niet tot het selecte gezelschap van harmonischen van de filterlengte behoort en in alle gevallen waarin géén sprake is van strikte periodiciteit introduceert het voortschrijdend gemiddelde distorties en kunstmatige fluctuaties (Kendall en Stuart, 1968, p. 377; Tintner, 1968, p. 50; Whittle, 1968, p. 64).

Waar het gaat om het effect op de *random* component is de werking van het voortschrijdend gemiddelde al evenmin loepzuiver. Het betreft hier een klassiek geval dat in de literatuur bekendstaat als het 'Slutzky effect', of het 'Slutzky-Yule effect' (Granger, 1964, p. 41-2; Jenkins en Watts, 1968, p. 297; Kendall en Stuart, 1968, p. 378 e. v.). Slutzky stelde vast dat successieve sommatie van *random* waarden leidde tot reeksen die een opmerkelijke tendens tot systematische oscillatie vertoonden (Slutzky, 1937, p. 105 e. v.). Dergelijke quasi-cyclische verschijnselen zijn artefacten die voortvloeien uit het feit dat successieve sommatie autocorrelatie tussen opeenvolgende waarden introduceert. Ook het voortschrijdend gemiddelde berust op successieve sommatie. Bijgevolg kan men stellen dat het bepalen van een voortschrijdend gemiddelde van een tijdreeks die een *random* component bevat, impliceert dat het Slutzky-Yule effect optreedt. De getransformeerde datareeksen bevatten zodoende naar alle waarschijnlijkheid een aantal artefacten waarvan noch het feitelijke verloop noch het relatieve gewicht bij voorbaat te bepalen is.

Op grond van het voorgaande kan men stellen dat de door Van Houwelingen c.s. uitgevoerde datatransformaties onder omstandigheden ernstige verontreiniging van de data tot gevolg kunnen hebben. In het geval van het gebruik van het 9-punts gemiddelde gaat het om distorties en onregelmatigheden die voortvloeien uit het effect van het voortschrijdend gemiddelde op de systematische component. Daarbij is het onder andere denkbaar dat een feitelijk aanwezige regelmatigheid dusdanig door de transformatie wordt verminkt dat hij als zodanig niet meer te herkennen valt. Bovendien bevat het residu een niet traceerbare dosis artefacten die een gevolg zijn van het Slutzky-Yule effect. In het geval van de eerste differenties zou men op grond van de klasse-verwantschap soortgelijke effecten verwachten. Door de omstandigheid echter dat het eerste differentiefilter zo 'kort' is en omdat weging met een negatieve factor plaatsvindt is de werking ervan anders. Dit houdt onder andere in dat de mate van distortie waarschijnlijk kleiner is. Daar staat tegenover dat differenties de eigenschap hebben om enerzijds systematische componenten te onderdrukken en anderzijds het op korte termijn alternerende deel van de reeks, in casu één aspect van de invloed van de *random* component, op te blazen<sup>6</sup>. In dit geval wordt niet alleen de invloed van de *random* component overdreven en de invloed van de systematische component onderschat maar ontstaat bovendien een artefact in de vorm van een tweejaars oscillatie (Yule, 1921, p. 515)<sup>7</sup>.

Door de aanwezigheid van dergelijke dataverontreinigingen is de betrouwbaarheid van uitspraken die aan een toets op de verkregen residuen worden ontleend aan gerede twijfel onderhevig.

#### 4. De toetssteen des aanstoots

Het voorgaande impliceert een kwalificatie bij voorbaat die echter als zodanig nog geen volledige verklaring vormt van het door Van Houwelingen c.s. verkregen resultaat.

De toets die de auteurs presenteren is in hoofdzaak gebaseerd op de correlatierekening. Bij de toepassing van een dergelijke techniek in de tijdreeksanalyse doet zich echter direct het probleem voor, dat men rekening moet houden met vertraagde reacties. Het is immers niet waarschijnlijk dat nieuwe werklozen onmiddellijk na ontvangst van hun ontslagbrief in de gracht springen of bezwijken aan een hartinfarct. In de regel zal sprake zijn van een

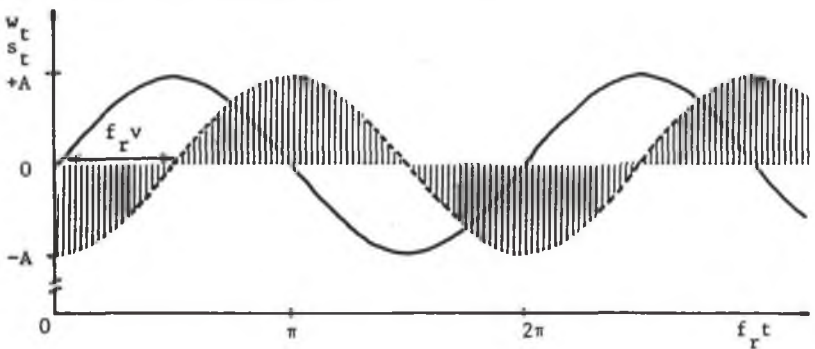
zekere vertragingstijd. Er is een *time-lag* tussen het moment waarop de impuls (van de verklarende variabele) optreedt en het moment waarop de reactie (van de te verklaren variabele) plaatsvindt.

De grootte van een dergelijke *time-lag* is niet a priori bekend. De auteurs proberen de daaraan inherente moeilijkheid te omzeilen door een enigszins experimentele benadering te kiezen. Zij berekenen correlaties voor verschillende situaties waarbij per situatie een andere *time-lag* wordt verondersteld. De veronderstelde *time-lag* ( $k$ ) wordt gevarieerd binnen het interval  $k = 0,1 \dots 15$ . Hieruit resulteren 16 correlatiecoëfficiënten die getabelleerd worden in een zogenaamd 'stem and leaf'-display (een soort histogram van correlatiecoëfficiënten). Aan de hand van de vorm van dit display trachten de auteurs vervolgens te beoordelen of er al dan niet van een verband tussen de geanalyseerde grootheden sprake is. Naar hun mening is er slechts dan van een verband sprake als het 'stem and leaf'-display, zoals zij dit noemen, 'positieve uitbijters' te zien geeft, dat wil zeggen als de berekende correlaties in het positieve deel van de display geconcentreerd zijn. De door hen geconstrueerde 'stem and leaf'-displays blijken vrij symmetrisch te zijn terwijl de grootste positieve en negatieve correlaties in dezelfde orde van grootte liggen. Op grond daarvan concluderen de auteurs dat er geen sprake is van significante verbanden tussen de geanalyseerde grootheden (t.a.p. p. 12 e.v.).

In hun streven een eenvoudige toets te ontwerpen begaan Van Houwelingen c.s. een noodlottige vergissing. Zij overzien een verraderlijke valkuil die onder het ogenschijnlijk onschuldige uiterlijk van hun methode verscholen ligt. De experimentele benadering waarbij de veronderstelde *time-lag* over een bepaald interval wordt gevarieerd is slechts consistent met het 'positieve uitbijter'-criterium als aan zeer specifieke voorwaarden is voldaan.

Om niet te vervallen in een al te technisch argument zal ik proberen om de aard van deze voorwaarden te illustreren aan de hand van een voorbeeld. Laten we aannemen dat het verloop van de verklarende variabele ( $w_t$ ) en de te verklaren variabele ( $s_t$ ) is als weergegeven in figuur 1<sup>8</sup>. Het valt in te zien dat er een strikt verband tussen beide reeksen bestaat. De te verklaren variabele volgt exact het spoor van de te verklaren variabele met dien verstande dat er sprake is van een *time-lag* van  $f_r v$ . Zou men reeks  $w_t$  over een afstand van  $f_r v$  naar rechts verschuiven, dan zou deze exact met reeks  $s_t$  samenvallen.

Figuur 1: Deel van het verloop van  $w_t$  en  $s_t$  voor  $f_r v = 0.5\pi$



In termen van figuur 1 komt het eerder genoemde variëren van de veronderstelde *time-lag* neer op het 'verschuiven' van  $w_t$  naar rechts. Vervolgens worden correlaties tussen de 'vershoven' reeks  $w_t$  en de (niet 'vershoven') reeks  $s_t$  berekend.

In de situatie die in figuur 1 is aangegeven, is de correlatie van  $w_t$  en  $s_t$  gelijk aan 0 (dit is evident omdat in die situatie de positieve deviaties van  $w_t$  even vaak samenvallen met positieve als met negatieve deviaties van  $s_t$ ). 'Verschuiven' we nu  $w_t$  een klein stukje naar rechts dan resulteert dit in een positieve correlatiecoëfficiënt (het aantal deviaties van  $w_t$  dat hetzelfde teken heeft als die van  $s_t$  neemt immers relatief toe). Herhaling van deze handeling levert steeds grotere correlatiecoëfficiënten op. De hoogste waarde treedt op als de 'verschuiving' gelijk is aan  $f_v$ . Dan vallen beide reeksen immers exact samen, hetgeen een correlatiecoëfficiënt van + 1 impliceert. Voorbij dit punt blijven de correlaties weliswaar positief maar daalt hun waarde geleidelijk weer tot 0 bij een 'verschuiving' van  $2 \times f_v$ . Tot nu toe was uitsluitend sprake van correlatiecoëfficiënten die of 0 of positief waren. In termen van een 'stem and leaf'-display zal dus een 'positieve uitbijter' ontstaan als de veronderstelde *time-lag* gevarieerd wordt tussen 0 en  $2 \times f_v$ .

Echt interessant wordt het pas wanneer men  $w_t$  verder dan  $2 \times f_v$  naar rechts vershuift. Dan worden de correlatiecoëfficiënten namelijk negatief! Hun absolute waarde neemt toe tot de 'verschuiving'  $3 \times f_v$  bedraagt. Op dit punt lopen beide reeksen exact tegengesteld aan elkaar, hetgeen correspondeert met een correlatiecoëfficiënt van -1. Voorbij dit punt dalen de absolute waarden weer en wordt het 0-niveau opnieuw bereikt bij een 'verschuiving' ter grootte van  $4 \times f_v$ . Bij variatie van de veronderstelde *time-lag* tussen  $2 \times f_v$  en  $4 \times f_v$  komt het bestaande verband tussen  $w_t$  en  $s_t$  zodoende tot uitdrukking in een 'negatieve uitbijter' in het 'stem and leaf'-display. Bij een variatie tussen 0 en  $4 \times f_v$  zijn er geen 'uitbijters' en is het *display* volkomen symmetrisch.

Deze kleine exercitie leidt tot de conclusie dat, gegeven het strikte verband tussen  $w_t$  en  $s_t$ , de vorm van het 'stem and leaf'-display volkomen wordt bepaald door het interval waarover men de veronderstelde *time-lag* varieert, in casu door de mate waarin men de reeksen ten opzichte van elkaar 'vershuift'. Anders geformuleerd: Het verband tussen  $w_t$  en  $s_t$  kan tot uitdrukking komen in verschillende vormen van het 'stem and leaf'-display al naar gelang men de veronderstelde *time-lag* binnen een ander interval varieert. Men mag zodoende pas een specifieke vorm van het stem and leaf-display als toetsingscriterium gebruiken als van tevoren vaststaat dat de veronderstelde *time-lag* over het daarmee corresponderende interval wordt gevarieerd. In het bovenstaande voorbeeld zou het 'positieve uitbijter'-criterium zodoende slechts consistent zijn met variaties van de veronderstelde *time-lag* tussen 0 en  $2 \times f_v$ . Variatie tussen 0 en  $4 \times f_v$  of tussen  $2 \times f_v$  en  $4 \times f_v$  zou daarentegen een contradictie opleveren.

De grenzen waarbinnen sprake is van consistentie worden volledig bepaald door de periodieke structuur van de geanalyseerde reeksen (in casu de 'periode' van het systematische patroon) en de grootte van de *feitelijke time-lag*<sup>9</sup>. Hanteert men een 'uitbijter' als criterium, dan dient tenminste<sup>10</sup> aan de voorwaarde voldaan te zijn dat het interval waarbinnen men de veronderstelde *time-lag* varieert kleiner of gelijk is aan de helft van de 'periode' van het systematische patroon. In het geval van Van Houwelingen c.s. is niet aan deze minimumvoorwaarde voldaan. Gezien de impliciete vooronderstellingen bij de datatransformatie gaan de auteurs er immers van uit dat de 'periode' van het systematische patroon korter is dan 9 jaar (zie paragraaf 2 boven). De veronderstelde *time-lag* daarentegen variëren zij over een interval van 15 jaar (t.a.p. p. 12). Dit kan slechts tot de conclusie leiden dat hun toetsingscriterium strijdig is met de eigenschappen van hun methode.

Om te illustreren hoe deze strijdigheid in de praktijk uitpakt heb ik een 'stem and leaf'-display à la Van Houwelingen c.s. berekend. Daarbij ben ik ervan uitgegaan dat de periode van het in figuur 1 weergegeven systematische patroon 4 jaar bedraagt. Dit komt ongeveer overeen met de lengte van de eerder genoemde 'minor business cycle'. De correlaties zijn evenals bij Van Houwelingen c.s. berekend over een interval van 0 tot en met

15 jaar. Het verband tussen  $w_t$  en  $s_t$  komt in dit geval tot uitdrukking in een volkomen symmetrische vorm van het 'stem and leaf'-display.

'Stem and leaf' – display voor  $w_{(t-k)}$  en  $s_t$ , met  $k=0,1 \dots 15$  bij een 4-jaars cyclus

-1 . 3, 7, 11, 15

0 . 0, 2, 4, 6, 8, 10, 12, 14

+1 . 1, 5, 9, 13

Het gaat hier uiteraard slechts om een gestileerd voorbeeld<sup>11</sup>. Het is echter 'realistisch' in die zin dat het bepaald niet is uitgesloten dat een soortgelijke situatie zich voordoet in de analyse die Van Houwelingen c.s. presenteren. Hun vergissing berust in een dergelijk geval op het feit dat onder de door henzelf gedefinieerde omstandigheden het bestaan van een verband tot uitdrukking komt in een symmetrische vorm van het 'stem and leaf'-display, terwijl zij juist op grond van deze symmetrie zouden concluderen dat er niet van een significant verband sprake kan zijn.

De protagonist van Harvey Brenners' hypothese die na deze constatering opgelucht ademhaalt en triomfantelijk concludeert dat de 'stem and leaf'-displays die Van Houwelingen c.s. presenteren, vrijwel symmetrisch zijn en dat *dus* sprake is van een verband tussen werkloosheid en sterfte, moet ik helaas ook teleurstellen. In de eerste plaats omdat rekening houdend met de eerder genoemde dataverontreiniging de symmetrie mogelijk verklaard wordt door een schijnverband dat eventueel uit het Slutsky-Yule effect kan ontstaan. In de tweede plaats omdat onder de door Van Houwelingen c.s. aangegeven verhoudingen geen enkele specifieke vorm van het 'stem and leaf'-display eenduidig interpreteerbaar is. De bovengenoemde symmetrie bijvoorbeeld treedt in die vorm niet in alle gevallen op. Er kan ook sprake zijn van partiële asymmetrie<sup>12</sup>.

In principe kan een verband zodoende in allerlei vormen van het 'stem and leaf'-display tot uitdrukking komen. Elke concrete vorm correspondeert met een bepaalde combinatie van de feitelijke periodelengte, de feitelijke *time-lag*, de veronderstelde *time-lag* en het interval waarbinnen de veronderstelde *time-lag* wordt gevarieerd. Daarom is het pas mogelijk om de vorm van het 'stem and leaf'-display als toetsingscriterium te gebruiken indien men over exacte a priori kennis beschikt van zowel de 'periode' als van de *time-lag* waardoor het verband tussen systematische componenten wordt gekarakteriseerd. Zolang deze a priori kennis ontbreekt is het onmogelijk om betrouwbare uitspraken aan de vorm van de displays te ontleen. In de omstandigheden die Van Houwelingen c.s. definiëren neemt deze onmogelijkheid een specifieke gedaante aan. Vanwege het feit dat in hun definitie de maximale veronderstelde *time-lag* groter is dan de maximaal toegestane 'periode' lengte van de systematische component, leidt hun operationalisatie tot foutieve conclusies in geval er sprake is van een systematisch verband tussen de geanalyseerde variabelen. Zij zullen het bestaan van een dergelijk verband immers steeds ontkennen, hetzij bij een symmetrisch 'stem and leaf'-display op grond van de symmetrie zelve (t. a. p. 12), hetzij bij een partieel asymmetrisch display op grond van de gelijkheid van de grootste positieve en de grootste negatieve correlaties (idem).

## 5. Besluit

Uit het voorgaande moge blijken dat de resultaten die Van Houwelingen c.s. boeken met ernstige twijfel omgeven zijn. De betrouwbaarheid van hun uitspraken wordt al bij voorbaat gerelativeerd door de effecten van de datatransformaties die zij uitvoeren. Zelfs op basis van de gunstige veronderstelling dat hun filters perfect werken, bestaat de mogelijkheid dat ze de hypothese ten onrechte verwerpen, hetzij doordat een verband dat behoort bij een systematische component met een lange periode bij voorbaat wordt weggefilterd, hetzij doordat een relatief sterke *random* component volledig intact gelaten wordt waardoor sporen van een verband naar de achtergrond verdwijnen. Bij de realistische veronderstelling dat de filters niet perfect werken, zijn fouten van verschillende orde denkbaar. De hypothese kan ten onrechte worden verworpen wanneer filterdistortie de sporen van een verband dermate verminkt dat het als zodanig niet meer herkenbaar is. Omgekeerd is het denkbaar dat de hypothese ten onrechte niet verworpen wordt doordat op basis van het Slutsky-Yule effect schijnrelaties uit de datatransformatie ontstaan.

Ondanks alle relativeringen vooraf is het uiteindelijk de kwaliteit van het toetsingsinstrument dat de doorslag moet geven. Helaas is het juist de toetssteen die een steen des aanstoots vormt. Men kan op eenvoudige wijze aantonen dat het toetsingscriterium dat Van Houwelingen c.s. hanteren strijdig is met de eigenschappen van de methode die ze toepassen. Daardoor vertoont hun toets onder de door henzelf bepaalde omstandigheden de chronische neiging om het bestaan van een verband tussen de geanalyseerde grootheden te ontkennen. Bezien vanuit dit standpunt bevat hun conclusie dat er niet van een significant verband tussen werkloosheid en gezondheid respectievelijk sterfte sprake is, dan ook geen enkele verrassing meer.

De dataverontreiniging en de defecten van het toetsingsinstrument brengen met zich dat de resultaten niet eenduidig interpreteerbaar zijn. Degenen die definitieve zekerheid willen hebben omtrent de juistheid of onjuistheid van de conclusie van Van Houwelingen c.s., zijn zodoende op nieuw onderzoek aangewezen.

Het is daarbij aan te bevelen om allereerst te expliciteren wat men precies onder een verband tussen tijdreeksen verstaat en welke rol men daarbij aan *time-lags* toekent. Vervolgens zou men de periodieke structuur van de relevante tijdreeksen, inclusief de betrokken *time-lags*, aan een nauwgezet empirisch onderzoek moeten onderwerpen. Op grond van de daarbij verworven inzichten zou men dan aan toetsing toe zijn.

Bij een dergelijk onderzoek lijkt de traditionele 'tijdsdomein' correlatietechniek die Van Houwelingen c.s. toepassen weinig perspectief te bieden. Het ligt daarom meer voor de hand om zogenaamde 'frequentie-domein' technieken als de Fourier- en Spectraal-analyse (Granger, 1964; Jenkins en Watts, 1968) toe te passen<sup>13</sup>. Behalve een krachtig instrument voor de analyse van tijdreeksen bieden dergelijke technieken mogelijkheden tot rechtstreekse toetsing van verbanden binnen het frequentiedomein door middel van de zogenaamde coherentiespectra (Granger, 1964, m.n. hfdst. 5 en 6). Men dient er zich echter van bewust te zijn dat de frequentie-domein techniek geen panacee is die moeite- en probleemloos kan worden toegepast. Aangezien het hier volgens Van Houwelingen c.s. echter gaat om een klassiek probleem voor de sociale wetenschappen (t.a.p. p. 6) is het altijd de moeite van het proberen waard.

## Noten

1. In de realiteit zal zelden of nooit sprake zijn van strikte periodiciteit. Het gaat veelal om fluctuaties waarop de term periode in strikte zin niet van toepassing is. Er is dan sprake van



een verdeling van 'periodelengten' rond een 'quasi-periode' waarvan de 'lengte' het gemiddelde van de afstand tussen extremen is. Men zou daarom consequent de uitdrukking 'periode dan wel quasi-periode' moeten gebruiken. Dit zou tot nogal omstandige formuleringen leiden. Daarom zet ik in het vervolg de term 'periode' tussen aanhalingstekens om aan te geven dat ik 'periode dan wel quasi-periode' bedoel.

2. Met uitzondering van de middenterm en de tekens van de termen heeft de reeks van gewichten dezelfde gedaante als die behorende bij een systematische component waarvan de vorm een tweede- of derdegraads polynoom is. De middenterm en de tekens van de termen zijn echter niet consistent met een deelterm ter grootte van 231 (t. a. p. p. 21). Ik neem aan dat hier van een drukfout sprake is en ga er verder van uit dat de auteurs een wegingsvector conform Kendall en Stuart (Kendall en Stuart, 1968, p. 368) voor ogen hebben, te weten:  $(-21, 14, 39, 54, 59, 54, 39, 14, -21)/231$ . *Mocht het zo zijn dat niet van een drukfout sprake is en dat inderdaad smoothing heeft plaatsgevonden door middel van de aangegeven tegenstrijdige hybride, dan is het aan de auteurs om een ander recht te zetten.*

3. De toevoeging 'analoog' dient ter onderscheiding van 'digitaal'. Bij het digitaal filteren gaat men uit van informatie uit het 'frequentiedomein' terwijl analoge filtering op grond van 'tijdsdomein'-informatie plaatsvindt; zie bijvoorbeeld Gold en Rader (1969).

4. Dat wil zeggen: behoudens het later te behandelen Slutsky-Yule effect.

5. Hogere harmonischen zijn de fracties  $P/m+1$ , waarbij  $P$  = periode en (integer)  $m = 1, 2, \dots, n$ . Lagere harmonischen zijn de veelvouden  $(m+1)P$ , waarin (integer)  $m = 1, 2, \dots, n$ .

6. Op dit effect berust de werking van de eerder genoemde 'variate difference' methode. Door de *random* component over een aantal ronden 'op te blazen' elimineert men de invloed van de systematische component en is men in staat de variantie van de *random* component indirect te schatten.

7. Volgens Yules studie wordt niet slechts de *random* component geaccentueerd maar wordt tevens een artefact in de vorm van de twee-jaars oscillatie gecreëerd (Yule, 1921, p. 515). Dit resultaat levert mogelijk een interessante verklaring op voor het feit dat Van Houwelingen c.s. zo'n 'merkwaardig' korte *time-lag* tussen werkloosheid en sterfte aan levercirrose vinden (t. a. p. p. 13). Volgens Yule is de verwachte autocorrelatie tussen eerste differenties van de *random*reeks  $-0.5$  voor opeenvolgende items en  $0.0$  voor andere (positieve) waarden van de 'lag' (Yule, 1921, p. 520). Het is dus zeer wel denkbaar dat de auteurs een artefact in plaats van een 'merkwaardige' samenhang op het spoor zijn.

8. Eenvoudigheidshalve is aangenomen dat de verklarende variabele  $w_t$  een cyclisch patroon volgt:  $w_t = A \sin(f_r t)$ , waarbij:  $A$  = amplitude,  $f_r$  = frequentie (in radialen per tijdseenheid) en  $t$  = tijdseenheid. Verder is aangenomen dat er een strikt verband is met de te verklaren variabele  $s_t$  die het verloop van  $w_t$  exact weerspiegelt met dien verstande dat er sprake is van een *time-lag* ter grootte van  $v$  tijdseenheden, zodat:  $s_t = \sin(f_r(t+v))$ .

9. In het voorbeeld is de periode van het systematisch patroon gelijk aan  $2\pi$  en de feitelijke *time-lag* is gelijk aan  $0.5\pi$ . Noemen we  $f_r k$  ( $k=0, 1, 2, \dots, m$  tijdseenheden) de veronderstelde *time-lag*, dan zijn de voorwaarden waaronder het 'stem and leaf'-display strikt positieve 'uitbijters' bevat als volgt weer te geven:  $f_r v - 0.5\pi \leq f_r k \leq f_r v + 0.5\pi$  ( $f_r v \geq 0$ ).

10. Hierbij wordt eenvoudigheidshalve afgezien van de voorwaarden die gesteld zijn aan de verhouding tussen het interval waarbinnen de veronderstelde *time-lag* wordt gevarieerd en de feitelijke *time-lag*.

11. Mijn argument is gebaseerd op de aanname dat de residuen een cyclisch patroon volgen. Men zou hieruit ten onrechte de conclusie kunnen trekken dat het verschil van mening tussen Van Houwelingen c.s. en mijzelf slechts berust op een verschil in assumpties betreffende de structuur van de tijdreeksen (cyclisch versus *random*). Het gaat echter om de structuur van de *residuen*. Door de datatransformaties wordt ook een reeks *random* waar-

den omgezet in een quasi-cyclisch residu vanwege het Slutsky-Yule effect. Mijn argument blijft ook dan onverkort gehandhaafd.

12. Wanneer de grootste veronderstelde *time-lag* geen geheeltallig veelvoud is van de periode van de systematische component, treedt gedeeltelijke asymmetrie op. In een dergelijk geval kan (uiteraard afhankelijk van de lengte van het feitelijke *time-lag* ten opzichte van de kleinste veronderstelde *time-lag*) het aantal positieve correlaties groter zijn dan het aantal negatieve of omgekeerd. Wat wel steeds blijft gelden is dat de grootste positieve en de grootste negatieve correlaties gelijk zijn. Om deze reden duid ik de laatste situatie dan ook als 'partieel asymmetrisch' aan.

13. Ik bedoel hiermee uiteraard de toepassing van de Fourier- of de Spectraal-analyse als analytisch en statistisch instrument en niet als een opstapje voor een louter illustratieve toepassing zoals bij Brenner het geval is (Brenner, 1971).

## Geraadpleegde literatuur

- Brenner, M. H., 'Economic Changes and Heart Disease Mortality', *American Journal of Public Health*, 61 (1971), 606-611.
- Brenner, M. H., 'Trends in Alcohol Consumption and Associated Illness. Some Effects of Economic Changes', *The American Journal of Public Health*, 65 (1975), 1279-1292.
- Brenner, M. H., 'Health Costs and Benefits of Economic Policy', *International Journal of Health Services*, 7 (1977), 581-623.
- Brenner, M. H., 'Mortality and the National Economy. A Review, and the Experience of England and Wales, 1936-1976', *The Lancet*, 15 (1979), 568-573.
- Brenner, M. H., 'Economic Change and Mortality in Post-War France. An Economic Model' (unpublished article, 1983).
- Duijn, J. van, *The Long-Wave in Economic Life*, London 1983.
- Gold, B., en C. Rader, 'Digital Processing of Signals'. New York 1969.
- Granger, C. W. J., en M. Hatanaka, *Spectral Analysis of Economic Time Series*, Princeton 1964.
- Jenkins, G. en Watts, D., *Spectral Analysis and its Applications*, San Fransico 1968.
- Kendall, M. G., en A. Stuart, *The Advanced Theory of Statistics*, Vol. III (Design and Analysis and Time Series), 2e ed., London 1968.
- Houwelingen, J. van, F. Tazelaar en A. Verbeek, 'Werkloosheid, gezondheid en sterfte in naoorlogs Nederland'. *Sociologische Gids*, 31 (1984), 2-23.
- Melnik, M., *Long Fluctuations in Real Series of American Economy*, Bureau of Economic and Business Research, Kent 1969.
- Reijnders, J., 'Perspectivistic Distortion. A note on some fundamental problems concerning the approximation of trends and trend-cycles', *Social Science Information*, 23 (1984), 411-426.
- Slutsky, E. E., 'The Summation of Random Causes as the Source of Cyclic Processes'. *Econometrica*, 5 (1937), 105-146.
- Tintner, G., *The Variate Difference Method*, Bloomington Ind., 1940.
- Tintner, G., 'Time Series'. *International Encyclopedia of the Social Sciences*, 16 (1968), 47-59.
- Tukey, J. W., *Exploratory Data Analysis*, Addison Wesley (Mass.), 1977.
- Whittle, P., 'Time Series. Advanced Problems'. *International Encyclopedia of the Social Sciences*, 16 (1968), 59-70.
- Yule, G. Udny, 'On the Time-correlation Problem, with Especial Reference to the Variate-difference Correlation Method', *Journal of the Royal Statistical Society*, 34 (1921), 497-526.