

Meting van beleidseffecten met tijdreeksanalyse

H. Folmer^{*}

Samenvatting

In dit artikel wordt een meetmethode van beleidseffecten beschreven, welke toegepast kan worden in situaties waarin informatie ontbreekt met betrekking tot andere belangrijke verklarende variabelen dan de beleidsvariabelen, zodanig dat specificatiefouten vermeden worden. Voorwaarde voor toepassing is dat twee tijdreeksen beschikbaar zijn: een pre-interventietijdreeks van de doelvariabele en een multivariatie tijdreeks van de doelvariabele en de beleidsvariabele, welke betrekking heeft op de periode waarin de beleidsvariabele werkzaam is.

De eerste tijdreeks kan veelal beschreven worden met een "multiplicative seasonal autoregressive integrated moving average model" (SARIMA model), waarin de effecten van allerlei relevante verklarende variabelen impliciet zijn opgenomen. Indien de relaties tussen de doelvariabele en de verklarende variabelen in de tweede periode dezelfde zijn als gedurende de eerste periode, zijn de veranderingen in het tijdspatroon van de doelvariabele het gevolg van effecten van de beleidsvariabele. Deze effecten kunnen onder bepaalde omstandigheden zuiver geschat worden door de tweede tijdreeks te specificeren als functie van het op basis van de eerste tijdreeks geschatte SARIMA model en de beleidsvariabele.

Deze methode is toegepast om de effecten van de extra werkgelegenheidsprogramma's op de werkloosheid in de bouw in de provincie Groningen gedurende de periode 1972-1976 te schatten.

^{*} De auteur is verbonden aan de Economische Faculteit RUG, Postbus 800, 9700 AV Groningen. Tel. 050 - 116830.

1. Inleiding

Meting van de effecten van overheidsbeleid heeft zich reeds lange tijd in een grote belangstelling mogen verheugen, zowel van de zijde van politieke en bestuurlijke instanties, als van het wetenschappelijk onderzoek. Immers, inzicht in de effecten is een noodzakelijke voorwaarde voor bijsturing van het bestaande beleid en voor het ontwerpen van een efficiënt toekomstig beleid.

Dit artikel beschrijft een meetmethode die berust op de volgende gedachte. Wanneer beleidsinterventies effect hebben gehad op een bepaalde variabele, waarin men geïnteresseerd is (in het vervolg kortweg aan te duiden als "doelvariabele"), dan komt dit tot uitdrukking in het tijdreekspatroon van de doelvariabele. Er zijn meerdere varianten van meetprocedures mogelijk, die op dit uitgangspunt steunen. In dit artikel zal aandacht geschonken worden aan die variant, die geschikt is voor situaties waarin informatie ontbreekt met betrekking tot factoren die naast de beleidsvariabelen van grote invloed zijn op de doelvariabele. In het algemeen is het niet mogelijk in een dergelijke situatie zuivere schatters te verkrijgen van het effect van de beleidsvariabele op de doelvariabele (zie bijvoorbeeld Theil, 1971). De hier te beschrijven methode kan een oplossing bieden voor dit probleem.

De methode kan toegepast worden in een veelheid van situaties. In dit artikel zal als toepassing besproken worden de meting van de effecten van extra werkgelegenheidsprogramma's op de werkgelegenheid in de bouw in de provincie Groningen gedurende de periode 1972-1976.

2. Het meetmodel

We zullen de basistructuur van het meetmodel aan de hand van de meest eenvoudige situatie, waarbij er sprake is van één doelvariabele en één beleidsvariabele, uiteenzetten.

Stel dat de eerste van een aantal beleidsinterventies, plaatsvindt op tijdstip T . We nemen verder aan dat we met betrekking tot de doelvariabele y over twee tijdreeksen beschikken: $\{y_{r,m}^{(1)}\}$ voor de periode vóór tijdstip T en $\{y_{r,m}^{(2)}\}$ voor de periode na tijdstip T .

De indices r en m zijn tijdsindices, bijvoorbeeld jaar r en week m .

De periode tussen twee opeenvolgende waarnemingen wordt constant verondersteld.

Zoals bovengezegd, zal de reeks $\{y_{r,m}^{(2)}\}$ verschillen van de eerste reeks als het beleid effect heeft gehad. Om deze verschillen te kunnen signaleren is het nodig modellen voor beide reeksen te vinden.

2.1. Model voor de pré-interventie tijdreeks

De reeks $\{y_{r,m}^{(1)}\}$ wordt bepaald door allerlei variabelen, die geen beleidsvariabelen zijn. Deze variabelen, in het vervolg aan te duiden als autonome variabelen, representeren we in het volgende model:

$$y_{r,m}^{(1)} = \eta_{r,m} \quad (1)$$

In een veelvoud van gevallen kan $\eta_{r,m}$ gemodelleerd worden als een zogenaamd "multiplicatieve seasonal autoregressive integrated moving average model" oftewel SARIMA model, zoals ontwikkeld door met name Box en Jenkins (1976). Een dergelijk model heeft de volgende vorm:

$$\psi_{p_1}(B) \omega_{p_2}(B^S)(1-B^d)(1-B^S)^D \eta_{r,m} = \theta_{q_1}(B) \lambda_{q_2}(B^S) a_{r,m} \quad (2)$$

In dit model is $\{a_{r,m}\}$ een reeks onafhankelijke normaal verdeelde variabelen met gemiddelde gelijk aan nul en variantie σ_a^2 . Een dergelijke reeks wordt veelal "witte ruis" genoemd. B stelt de backshift operator voor met betrekking tot de tijdsindex r . Dat wil zeggen, dat $B \eta_{r,m}^h = \eta_{r-h,m}$. Evenzo is B^S de seizoenen backshift operator met periode s : $(B^S)^p \eta_{r,m} = \eta_{r,m-ps}$. Opgemerkt zij dat "seizoen" hier in ruime zin begrepen moet worden en ook betrekking kan hebben op, bijvoorbeeld, zich wekelijks herhalende verschijnselen.

De term $(1-B)^d$ stelt de trend van de doelvariabele voor over de tijdsindex r . Deze trend is het gevolg van allerlei belangrijke deterministische factoren, die niet expliciet in het model zijn opgenomen. Op dezelfde wijze representeert de term $(1-B^S)^D$ seizoenenstructuren.

De term $\psi_{p_1}(B)$ is een polynoom van de graad p_1 in de backshift operator B :

$$\psi_{p_1}(B) = \psi_1 B - \psi_2 B^2 - \dots - \psi_{p_1} B^{p_1}.$$

Evenzo is $\omega_{p_2}(B^S)$ een polynoom van de graad sp_2 .

Gezamenlijk beschrijven beide polynomen de doelvariabele, gecorrigeerd voor trend en seizoensfluctuaties met behulp van $(1-B)^d$ en $(1-B^s)^D$, als een functie van zichzelf in voorgaande periods. De termen $\Theta_{q_1}(B)$ en $\lambda_{q_2}(B^s)$ zijn polynomen van de graad q_1 , respectievelijk q_2 , en relateren de toevalsvariabelen $a_{r,m}$ aan de doelvariabele.

Het is voor het volgende gemakkelijker één tijdsindex, zeg t , te gebruiken. De dubbele index kan gemakkelijk omgezet worden in een enkele door de kleinste tijdseenheid als teleenheid te gebruiken.

Vaststelling van empirische waarden van de parameters van model (2) gebeurt in een iteratief proces van modelidentificatie, schatten van het geïdentificeerde model en controle op het geschatte model. De eerste parameter die geïdentificeerd moet worden is de periode s . Voorkennis en visuele inspectie van een plot van de ruwe data geven veelal belangrijke aanwijzingen over de waarde van s . Bovendien geven de geschatte autocorrelaties r_k en de geschatte partiële autocorrelaties $\hat{\phi}_{kk}$ informatie over s .

De geschatte autocorrelatie r_k is gedefinieerd als

$$r_k = \frac{1}{N} \sum_{i=1}^{N-k} (y_i - \bar{y})(y_{i+k} - \bar{y}), \quad k = 0, 1, 2, \dots, K \quad (3)$$

waarbij N het aantal waarnemingen en $\bar{y}$ het gemiddelde van de te analyseren tijdreeks is.

De geschatte partiële autocorrelatie is gedefinieerd als

$$\hat{\phi}_{kk} = \frac{|R_k^*|}{|R_k|} \quad k = 1, 2, \dots, K \quad (4)$$

waarbij $| \cdot |$ de determinant is en R_k de autocorrelatiematrix:

$$R_k = \begin{bmatrix} 1 & r_1 & r_2 & \dots & r_{k-1} \\ r_1 & 1 & r_1 & \dots & r_{k-2} \\ r_2 & r_1 & 1 & \dots & r_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & r_{k-2} & r_{k-3} & \dots & 1 \end{bmatrix}$$

R_k^* is de autocorrelatiematrix met de laatste kolom vervangen door

$\{r_1, r_2, \dots, r_k\}^T$. De autocorrelaties en partiële autocorrelaties kunnen geschat worden met behulp van een door Pack (1974) ontwikkeld computerprogramma.

Dit programma berekent ook de variantie van r_k terwijl de variantie van $\hat{\phi}_{kk}$ bij benadering gelijk is aan $\frac{1}{n}$, waarbij n het aantal waarnemingen zonder restricties is. De schatters r_k en $\hat{\phi}_{kk}$ zijn bij benadering normaal verdeeld, zodat hypothesetests met betrekking tot deze parameters getoetst kunnen worden.

In geval de periode de waarde s' heeft zullen auto- en partiële autocorrelaties met vertragingen van waarden rondom $s'-1, 2s'-1, \dots$ significant van nul verschillen.

In de tweede stap van het identificatieproces worden de waarden van d en D bepaald. Dit kan op de meest eenvoudige wijze geschieden door de variantie van $(1-B)^d (1-B^s)^D y_t^{(1)}$ te berekenen voor verschillende waarden van d en D en die waarden te kiezen waarvoor de variantie minimaal is.

In de derde stap moeten p_1, q_1 en p_2, q_2 bepaald worden. Het eerste paar kan afgeleid worden van de geschatte autocorrelaties en partiële autocorrelaties met lage waarden voor k en het tweede paar van dezelfde grootheden met waarden van k rondom $s'-1, 2s'-1, \dots$. De afleiding voor beide paren kunnen afzonderlijk, maar op dezelfde wijze geschieden. Vandaar dat de hieronder te beschrijven identificatieprocedures op beide paren van toepassing is, zodat de in i es weggelaten kunnen worden.

Op de eerste plaats geldt dat $p+q$ in de praktijk bijna altijd kleiner dan of gelijk aan twee is. Ten tweede, de autocorrelaties van een puur autoregressief proces ($q=0$) gaan voor toenemende k geleidelijk naar nul, terwijl $\hat{\phi}_{kk}$ voor $k > p$ niet significant van nul verschilt.

De eigenschappen van een zuiver moving average proces zijn tegengesteld: de autocorrelaties zijn voor $k > p$ niet significant verschillend van nul, terwijl de partiële autocorrelaties met toenemende k geleidelijk naar nul dalen.

Bij een gemengd autoregressief moving average proces gaan zowel de autocorrelaties als de partiële autocorrelaties geleidelijk naar nul bij groter wordende k .

Opgemerkt zij dat de hier geschetste gedragingen van de autocorrelaties en partiële autocorrelaties ideaal-typisch zijn. In de praktijk kunnen door het toeval veroorzaakte afwijkingen voorkomen.

Wanneer het model geïdentificeerd is, wordt het geschat. Niet-lineaire kleinste kwadraten schatters worden meestal gebruikt. Deze schatters zijn bij beandering normaal verdeeld. Voor de niet-lineaire kleinste kwadraten schattingsprocedure kan wederom gebruik gemaakt worden van het ontwikkelde computerprogramma door Pack (1974). Dit programma berekent ook de covariantiematrix van de schatters.

Tenslotte moet het geschatte model op eventuele onvolkomenheden gecontroleerd worden. Hierbij moet onderscheid gemaakt worden tussen onvolkomenheden in de algemene structuur van het geschatte model, zoals bijvoorbeeld een te lage orde van een of meerdere polynomen, en de significantie van de individuele parameters.

Het genoemde computerprogramma berekent betrouwbaarheidsintervallen waaruit conclusies over significantie van de parameters getrokken kunnen worden. Onvolkomenheden in de algemene structuur van het model komen meestal tot uitdrukking in de geschatte residuen, welke eveneens door het computerprogramma berekend worden. De residuen van een correct geïdentificeerd en geschat model gedragen zich als witte ruis.

De hypothese dat een reeks witte ruis is kan getoetst worden met de χ^2 toets. De χ^2 waarde van een reeks wordt gegeven door Pack's (1974) computerprogramma.

Bestaat er aanleiding de hypothese dat de reeks residuen een realisatie van witte ruis is, te verwerpen dan moet het voorliggende model herzien worden. De geschatte residuen van het inadequate model kunnen een belangrijke rol spelen bij de herziening, hetgeen we illustreren aan de hand van het volgende voorbeeld.

Stel dat het in eerste instantie geschatte (onvolkomen) model luidt:

$$\psi_{p_1}^{(B)} \omega_{p_2}^{(B^S)} \eta_t = \theta_{q_1}^{(B)} a'_t \quad (5)$$

en dat we voor de residuen vinden

$$a'_t = \lambda_{q_2}^{(B^S)} u_t \quad (6)$$

met $\{u_t\}$ witte ruis

Substitutie van (6) in (5) geeft

$$\psi_{p_1}^{(B)} \omega_{p_2}^{(B^S)} \eta_t = \theta_{q_1}^{(B)} \lambda_{q_2}^{(B^S)} u_t \quad (7)$$

Vervolgens wordt model (7) geschat en gecontroleerd op onvolkomenheden. Eventueel wordt deze procedure herhaald.

2.2. Model voor de tijdreeks beïnvloed door de beleidsvariabele

De modelkeuze voor de tijdreeks, beïnvloed door het beleid, wordt in sterke mate bepaald door de aard van de beleidsinterventies. Globaal gesproken kan men twee soorten interventies onderscheiden: stapinterventies, welke een permanente verandering van een element van het economisch systeem betekenen, zoals de ontsluiting van een gebied door een nieuwe weg, en schokinterventies, welke een tijdelijke verandering van een onderdeel van het systeem betekenen, zoals een "tijdelijke" belastingverhoging.

Opgemerkt zij dat herhaalde interventies, zoals extra werkgelegenheidsprogramma's, veelal opgevat kunnen worden als een rij van schokinterventies.

Bij de keuze van een model voor de tijdreeks beïnvloed door één of meerdere beleidsinterventies gaan we uit van de volgende veronderstellingen. Op de eerste plaats nemen we aan dat de verschillen tussen de tijdreeksen $\{y_t^{(1)}\}$ en $\{y_t^{(2)}\}$ veroorzaakt worden door de beleidsvariabele. Dit betekent dat de relaties tussen de doelvariabele en de niet-beleidsvariabele na tijdstip T verondersteld worden dezelfde te zijn als vóór tijdstip T. Deze rigide voorwaarden kan afgezwakt worden door autonome variabelen, waarin veranderingen hebben plaats gevonden en waarover informatie beschikbaar is, in het model voor $y_t^{(2)}$ op dezelfde wijze op te nemen als de beleidsvariabele. Ten tweede veronderstellen we een lineair additieve modelstructuur. Ten derde nemen we aan dat de beleidsvariabele en de variabelen, die de pré-interventie tijdreeks bepalen, onafhankelijk zijn. Dit alles betekent, dat $\{y_t^{(2)}\}$ in algemene zin met het volgende model beschreven kan worden:

$$y_t^{(2)} = \eta_t + f(\xi_t) \quad (8)$$

waarbij ξ de beleidsvariabele is en $f(.)$ een zogenaamd transfer functie model is.

Bij de specificatie van het transfer functie model, dat de relatie beschrijft tussen de beleidsvariabele en de doelvariabele, spelen theoretische en a priori kennis met betrekking tot de vorm van en de

vertragingen in de effecten een belangrijke rol (voor een overzicht van een aantal mogelijke modellen zij verwezen naar Folmer, 1983). In het navolgende kiezen we het simpele, door Box en Jenkins (1976) ontwikkelde model:

$$f(\xi_t) = \frac{\epsilon(B)}{\delta(B)} \xi_{t-b} \quad (9)$$

waarbij

$$\epsilon(B) = \epsilon_0 - \epsilon_1 B - \dots - \epsilon_m B^m$$

$$\delta(B) = 1 - \delta_1 B - \dots - \delta_n B^n$$

b de tijdsspanne weergeeft welke verloopt tussen de periode van interventie en de periode waarin de beleidsvariabele het eerste effect heeft.

Indien een adequaat model voor $\{y_t^{(1)}\}$ geschat is, maakt de eerste van de bovenstaande veronderstellingen het mogelijk om η_t in (8) te vervangen door het geschatte model:

$$\hat{\Lambda} a_t$$

waarbij

$$\hat{\Lambda} = \frac{\hat{\theta}_{q_1}(B) \hat{\lambda}_{q_2}(B^{\hat{s}})}{\hat{\psi}_{p_1}(B) \hat{\omega}_{p_2}(B^{\hat{s}}) (1-B)^{\hat{d}} (1-B^{\hat{s}})^{\hat{D}}}$$

Dit geeft

$$y_t^{(2)} = \hat{\Lambda} a_t + \frac{\epsilon(B)}{\delta(B)} \xi_{t-b} \quad (10)$$

Opgemerkt zij, dat de effecten van de autonome variabele, vertegenwoordigd zijn in $\hat{\Lambda}$. Vandaar dat modelspecificatie (10) het mogelijk maakt het effect van de beleidsvariabele zuiver te schatten.

Vermenigvervuldigd van (10) met $\hat{\Lambda}^{-1}$ leidt tot een getransformeerde reeks $\{\tilde{y}_t^{(2)}, \tilde{\xi}_t\}$ welke beschreven kan worden als

$$\tilde{y}_t^{(2)} + \frac{\epsilon(B)}{\delta(B)} \tilde{\xi}_{t-b} = a_t \quad (11)$$

Vaststelling van de empirische waarden van model (11) geschiedt wederom in een iteratief proces van identificeren (het vaststellen van de ordes van de beide polynomen, en van de waarde van b),

schatten en controleren.

Identificatie van $\epsilon(B)$, $\delta(B)$ en van bin (11) wordt vergemakkelijkt door "prewhitening" van de beleidsvariabele. Dat wil zeggen dat een SARIMA model voor de beleidsvariabele wordt geschat, dat de beleidsvariabele transformeert tot witte ruis. Dezelfde transformatie wordt toegepast op de doelvariabele. Vervolgens wordt voor beide getransformeerde variabelen de crosscorrelatie functie berekend. Voor een definitie van deze functie zij verwezen naar Box en Jenkins (1976, hoofdstuk 11). Berekening kan geschieden met het door Pack (1974) ontwikkelde programma. Met behulp van de crosscorrelaties kunnen aanwijzingen verkregen worden over de waarde van b en de graden van $\epsilon(B)$ en $\delta(B)$. Voor verdere details zij verwezen naar Box en Jenkins (1976, i.h.b. de hoofdstukken 10 en 11).

Het reeds vele malen genoemde computerprogramma van Pack produceert ook niet-lineaire kleinste kwadraten schattingen van model (11) alsmede standaardfouten.

In de laatste fase dient de algemene structuur van het model op onvolkomenheden te worden onderzocht. Ook voor de onderhavige tijdreeks geldt dat de geschatte residuen van een correct model zich als realisatie van een toevalsvariabele gedragen. Bovendien dienen de geschatte parameters op significantie getoetst te worden. Zowel de geschatte residuen als betrouwbaarheidsintervallen voor de geschatte parameters worden door het genoemde computerprogramma berekend.

In het bovenstaande is herhaaldelijk gesproken over "toetsen van hypothesen". Opgemerkt zij, dat het begrip toetsen hier niet verstaan moet worden in de gebruikelijke zin van het woord, waarbij een gegeven hypothese volgens de regels van de statistische beslissingstheorie al of niet geaccepteerd wordt met een van te voren vastgestelde kans op een fout van de eerste soort. Immers, in de meeste van de hier beschreven situaties is een hypothese op voorhand niet aanwezig. Integendeel, de steekproef data, welke voor toetsingsdoeleinden worden gebruikt, zijn in de identificatie fase geanalyseerd om een acceptabel model op te sporen. Deze werkwijze, waarvoor fraaie uitdrukkingen als "data mining", "fishing", "grubbing" en zelfs "martelen van de data" geïntroduceerd zijn, maakt gebruik van alle typische eigenaardigheden van de ene voor handen zijnde steekproef. Vandaar dat de goodness of model fit met be-

trekking tot de beschikbare data waarschijnlijk groter is dan met betrekking tot de populatie. Het gevolg is dat de kans op een fout van de eerste soort gewoonlijk veel groter is dan de van te voren vastgestelde kans.

We zullen dit probleem hier negeren. Voor nadere informatie zij verwezen naar onder andere Lovell (1983).

3. Effecten van extra werkgelegenheidsprogramma's in de provincie Groningen

In deze sectie wordt de boven beschreven methode toegepast om de effecten van de beleidsvariabele extra werkgelegenheidsprogramma's te meten op de doelvariabele werkloosheid in de sector grond-, weg-, en waterwerken in de provincie Groningen.

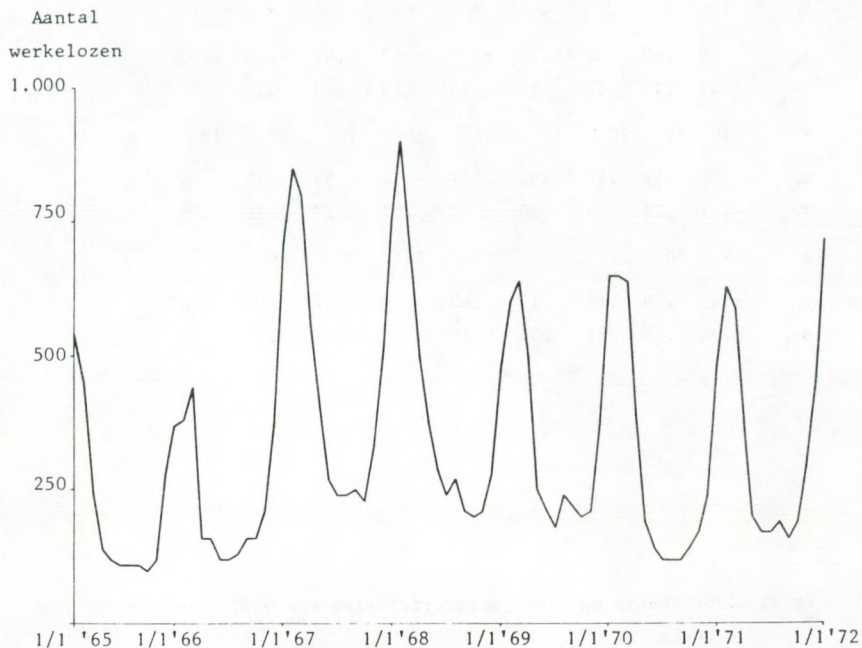
De doelvariabele is gemeten als het aantal officieel geregistreerde werkelozen. Informatie met betrekking tot de doelvariabele is beschikbaar in de vorm van maandelijks, door het CBS gepubliceerde, tijdreeksen voor de pré-interventieperiode 1965-1971 en voor de interventie periode 1972-1976.

Gedurende deze laatste periode werden de extra werkgelegenheidsprogramma's gerealiseerd. Voor de beleidsvariabele, gemeten in miljoenen gulden, zijn officiële gegevens slechts beschikbaar op jaarbasis en wel voor de periode 1972-1976 (Ministerie van Economische Zaken, 1977).

Dit geringe aantal observaties vormt een ernstige belemmering om een adequaat model te bepalen. Een andere belemmering voor het verkrijgen van een nauwkeurig beeld van de beleidseffecten wordt gevormd door het feit dat in de gepubliceerde gegevens ook zijn opgenomen ten behoeve van gebouwen voor, bijvoorbeeld, sport en culturele voorzieningen. Vandaar dat de onderstaande coëfficiënt van de beleidsvariabele met de nodige voorzichtigheid geïnterpreteerd moet worden. Het navolgende is daarom voornamelijk een illustratie van het voorgaande.

Conform paragraaf 2 is de eerste stap van de meetprocedure het vinden van een adequaat model voor de doelvariabele op grond van de pré-interventietijdreeks van 84 waarnemingen.

Deze waarnemingen zijn grafisch weergegeven in figuur 1.



Figuur 1. Geregistreeerde werklozen in de sector grond-, weg- en waterwerken in de provincie Groningen in de periode 1965-1971.

Inspectie van figuur 1 en a priori kennis over het sterke seizoens gebonden karakter van de werkeloosheid in de bouw geven aanleiding om s gelijk te stellen aan 12. Dit vermoeden wordt nog versterkt door de geschatte auto- en partiële autocorrelatie functies welke weergegeven zijn in de tabellen 1 en 2. Met name de waarde van r_{11} geeft ondersteuning aan de hypothese $s=12$.

Tabel 1. Geschatte autocorrelaties r_k en hun standaardfouten SE_k van de originele reeks.

k	1	2	3	4	5	6	7	8	9
r_k	.78	.40	.02	-.25	-.37	-.42	-.40	-.30	-.05
se_k	.11	.16	.17	.17	.18	.19	.20	.21	.21
k	10	11	12	13	14	15	16	17	18
r_k	.30	.59	.68	.52	.21	-.11	-.34	-.45	-.49
se_k	.21	.22	.24	.26	.27	.27	.27	.28	.29
k	19	20	21	22	23	24	25	26	27
r_k	-.48	-.39	-.18	.12	.40	.50	.27	.12	-.13
se_k	.30	.31	.31	.31	.31	.32	.33	.33	.33

Tabel 2. Geschatte partiële autocorrelaties $\hat{\phi}_{kk}$ van de originele reeks

k	1	2	3	4	5	6	7	8	9
$\hat{\phi}_{kk}$	.78	-.55	-.14	.00	-.09	-.23	-.05	.11	.33
k	10	11	12	13	14	15	16	17	18
$\hat{\phi}_{kk}$	.28	.11	.01	-.20	-.12	-.10	-.16	.01	-.06
k	19	20	21	22	23	24	25	26	27
$\hat{\phi}_{kk}$	-.12	-.19	-.06	.06	.04	-.05	-.07	.03	-.03

De bij $\hat{\phi}_{kk}$ behorende standaardfouten zijn bij benadering gelijk aan $N^{-\frac{1}{2}} = (84)^{-\frac{1}{2}} = 0.11$.

Vervolgens worden de parameters d en D bepaald door het berekenen van de standaardvariatiën van de met $(1-B)^d (1-B^{12})^D$ getransformeerde reeks voor verschillende waarden van D en d . De resultaten zijn weergegeven in tabel 3.

Tabel 3. Standaarddeviaties voor verschillende waarden van d en D .

D/d	0	1	2
0	2267.1	1397.2	1306.2
1	151.4	85.2	119.3
2	2541.3	1443.5	1921.1

Uit tabel 3 blijkt dat het minimum bereikt wordt voor $d = 1$ en $D = 1$.

De χ^2 van de reeks getransformeerd met $(1-B)(1-B^{12})$ is 34.1 bij 27 vrijheidsgraden. Het kansniveau, dat wil zeggen de kans om een dergelijke χ^2 waarde door toeval te krijgen indien H_0 , dat deze reeks witte ruis is, waar is, ligt tussen 10% en 25%. Dus de nulhypothese kan niet worden verworpen op het 10% niveau. De auto- en partiële autocorrelatie functies van de getransformeerde reeks zijn weergegeven in de tabellen 4 en 5.

Tabel 4. Geschatte autocorrelaties r_k en standaardfouten SE_k van de reeks getransofmeerd met $(1-B)(1-B^{12})$.

k	1	2	3	4	5	6	7	8	9	10	11
r_k	.01	-.04	-.05	-.11	-.02	.00	-.09	.03	.17	.22	-.01
SE_k	.12	.12	.12	.12	.12	.12	.12	.12	.12	.13	.13
k	12	13	14	15	16	17	18	19	20	21	22
r_k	-.36	-.11	-.05	.02	.12	.01	.01	.11	-.04	-.17	-.29
SE_k	.13	.14	.15	.15	.15	.15	.15	.15	.15	.15	.15
k	23	24	25	26	27						
r_k	.13	-.16	.13	.15	-.02						
SE_k	.16	.16	.16	.16	.17						

Tabel 5. Geschatte partiële autocorrelaties $\hat{\phi}_{kk}$ van de reeks getransformeerd met $(1-B)(1-B^{12})$.

k	1	2	3	4	5	6	7	8	9	10	11
$\hat{\phi}_{kk}$	.01	-.04	-.05	-.11	-.02	-.02	-.10	.02	.16	.22	.00
k	12	13	14	15	16	17	18	19	20	21	22
$\hat{\phi}_{kk}$	-.36	-.11	-.02	.02	.10	.03	.00	-.03	-.10	-.09	-.15
k	23	24	25	26	27						
$\hat{\phi}_{kk}$	.26	-.33	.26	-.01	-.08						

De bij $\hat{\phi}_{kk}$ behorende standaardfouten zijn bij benadering gelijk aan $(71)^{-\frac{1}{2}} = 0.12$.

Uit de tabellen 4 en 5 blijkt dat voor lage waarden van k r_k en $\hat{\phi}_{kk}$ niet significant zijn. Evenmin valt er een geleidelijk afnemende tendens te signaleren. Vandaar dat we concluderen dat het model geen algemene moving average of autoregressieve component bezit.

Wat het seizoensmodel betreft blijken $\hat{\phi}_{12,12}$, $\hat{\phi}_{23,23}$, $\hat{\phi}_{25,25}$ significant te zijn, waarbij $\hat{\phi}_{12,12}$ het grootst is. De partiële autocorrelaties vertonen dus een afnemende tendens, hetgeen een seasonal moving average model suggereert. Bij de bepaling van de orde blijkt allen r_{12} significant te zijn. We concluderen daarom tot een MA(1) seizoensmodel.

Niet lineaire kleinste quadaten schatting geeft het volgende model.

$$(1-B)(1-B^{12}) \eta_t = (1-0.65B^{12}) a_t \quad (12)$$

met -0.49 en -0.85 als boven en benedengrenzen van het 95% betrouwbaarheidsinterval.

De χ^2 waarde van de residuen van model (12) is gelijk aan 14.9 met 26 vrijheidsgraden. Het kansniveau van de hypothese, dat de reeks residuen witte ruis ligt tussen 90% en 95%. Vandaar dat we dit model accepteren.

De tweede stap ter bepaling van de effecten van de extra werkgelegenheidsprogramma's bestaat in de schatting van model (11). Daartoe wordt de interventiereeks van de doelvariabele getransformeerd met

$$\frac{(1-B)(1-B^{12})}{1-0.69B^{12}} \quad (13)$$

Omdat voor de extra werkgelegenheidsprogramma's alleen jaarlijks data

beschikbaar zijn, zijn de getransformeerde maandelijkse data van de doelvariabele gesommeerd tot jaarlijkse totalen. Vervolgens zijn voor de extra werkgelegenheidsprogramma's eerste ordeverschillen berekend volgens de niet-seizoenscomponent (1-B) van model (12). Tenslotte is het volgende model geschat (t waarde tussen haakjes):

$$\hat{y}_t^{(2)} = -8.9 \hat{\xi}_t + a_t \quad R^2 = 0.72 \quad (14)$$

(3.6)

Uit model (14) kan het volgende geconcludeerd worden. De coëfficiënt in dit model heeft het juiste teken omdat het realistisch is te veronderstellen dat de extra werkgelegenheidsprogramma's een negatieve invloed hebben op de werkeloosheid. Bovendien, een eenzijdige t-toets van het gebruikelijke niveau 0.05 bij 4 vrijheidsgraden geeft dat het effect van de extra werkgelegenheidsprogramma's significant van nul verschilt. Tenslotte kan worden afgeleid dat een beleidsimpuls van 1 miljoen gulden leidt tot een daling van de werkeloosheid met 8.9 personen. Bij deze laatste conclusie zij nogmaals gewezen op de onvolkomenheden in de data en op het beperkte aantal waarnemingen, welke de identificatie en schatting van een meer adequaat model, met bijvoorbeeld vertragingen in de afhankelijke en/of onafhankelijke variabele, verhindert.

Referenties

- Box, G.E.P. and G.M. Jenkins. (1979), Time series analysis, Forecasting and Control, San Fransisco et al: Holden-Day.
- Folmer, H. (1983), Measurement of effects of regional economic Policy. Some methodological aspects. (Te verschijnen)
- Lovell, M.C. (1983), Data mining. Review of Economics and Statistics, 65, 1-12.
- Ministerie van Economische Zaken (1979). Nota regionaal sociaal-economisch beleid 1977 t/m 1980. 's Gravenhage: Staatsuitgeverij.
- Pack, D.J. (1979), "Computerprograms for the analysis of univariate time series models and single input transfer function models using the methods of Box and Jenkins." Mimeographed Paper, The Ohio State University, Columbus, Ohio.
- Theil, H. (1971), Principles of Econometrics, New York: Wiley.