

TYDREEKSANALYSE EN FOUTEN IN BNP-CYFERS EEN THEORETISCHE KORREKTIE EN HERINTERPRETATIE

door

Eric GHYSELS (*)
Aspirant N.F.W.O.

Centrum voor Ekonometrie en Management Science
V.U.B.

J.L. Kellogg Graduate School of Management
Northwestern University

Summary on p. 495

1. INLEIDING

In een artikelenreeks heeft de auteur samen met J. Vuchelen (hierna G-V) de fouten in de maandelijkse reële BNP-indexcijfer van DULBEA bestudeerd. Fouten worden gedefinieerd als het verschil tussen voorlopige cijfers en definitieve publikaties. In deze reeks van publikaties werd ervan uitgegaan (vooral [3]) dat rationaliteit bij het berekenen van voorlopige cijfers kan geïdentificeerd worden met witte ruis karakter voor fouten. In deze tekst wordt aangetoond dat rationaliteit noch een voldoende noch een nodige voorwaarde is voor witte ruis fouten. Dit heeft een aantal gevolgen voor de interpretatie van de resultaten die werden bekomen. In de tweede paragraaf tonen wij aan waarom witte ruis niet volgt uit rationeel gedrag en in een laatste paragraaf worden

(*)Onze dank gaat naar J. Vuchelen voor zijn kritische opmerkingen.

de bestaande resultaten gekorrigeerd in het licht van die bevindingen. Tot slot wordt nog gesuggereerd dat de term "rationaliteit" die wij, en vele anderen tot nog toe, hebben gebruikt eigenlijk totaal verkeerd is in de huidige kontekst.

2. THEORETISCHE KORREKTIE

Rationaliteit van voorspellingen slaat op het feit dat alle informatie beschikbaar met betrekking tot de te voorspellen kwantiteit op het ogenblik van de voorspelling wordt gebruikt. Dit theoretisch begrip heeft als groot nadeel dat het moeilijk empirisch te testen valt. Hetgeen de protagonisten ter kennis hebben bij het uitvoeren van voorspellingen is om evidente redenen moeilijk te achterhalen. Het blijft dan ook gissen naar de informatieverzameling I_t , het formele concept dat al de informatie inhoudt beschikbaar op ogenblik t , die men moet aanwenden om rationaliteit te toetsen van bijvoorbeeld DULBEA bij publikaties van BNP-indices. De publikaties van de BNP-indexcijfers zijn voorspellingen in die zin dat zij worden berekend op een ogenblik dat slechts een fragmentair beeld van de industriële activiteit in het land beschikbaar is.

De eerste publikatie van een BNP-cijfer voor een bepaalde maand steunt immers op zeer partiële informatie. DULBEA publiceerde in het verleden ieder kwartaal de cijfers in een statistische bijlage aan dit tijdschrift. Wij gaan een specifiek voorbeeld aanwenden om onze beweringen aan te tonen. De redenering is evenwel algemeen toepasbaar, ook voor de huidige half-jaarlijkse publikaties, en het specifieke voorbeeld doet dus geen afbreuk aan de algemeen geldigheid ervan.

We beschouwen het eerste kwartaalnummer, in dat nummer werd meestal het indexcijfer voor november en december van het voorgaande jaar voor de eerste maal gepubliceerd alsook dat voor de maand januari. Formeel hebben we drie informatieverzamelingen I_N^t , I_D^t en I_J^t welke alle beschikbare informatie omvatten met betrekking

tot de industriële activiteit in de respektievelijke maanden op het ogenblik van de publikatie t (einde eerste kwartaal). Zodoende dat wanneer DULBEA "rationeel" handelt, wij als publikatie hebben $E(BNP_N | I_N^t); E(BNP_D | I_D^t)$ en $E(BNP_J | I_J^t)$. Fundamenteel is echter wat I_N^t, I_D^t en I_J^t omvatten. In G-V [2] werd aan de hand van de "informatiecurve" de stelling verdedigd dat $I_N^t > I_D^t > I_J^t$ waarbij ">" een orde van omvang weergeeft. Dit betekent dat DULBEA meer weet over activiteiten die plaatsgrepen in november dan die voor januari. Er is al meer statistisch materiaal beschikbaar, sommige bronnen zijn zelfs al aan herziening toe. De hypothese vond ook enige empirische verifikatie in de kwartaal seizoenstructuur die G-V [2] terugvinden in de geschatte modellen voor de eerste fout. Nemen we dus de vermelde orde als werkhypothese dan hebben we als fouten (onder hypothese van rationaliteit) :

$$\begin{aligned} BNP_N - E(BNP_N | I_N^t) &= \epsilon_N \\ BNP_D - E(BNP_D | I_D^t) &= \epsilon_D \\ BNP_J - E(BNP_J | I_J^t) &= \epsilon_J \end{aligned} \tag{2.1}$$

Alzo vormen we een tijdreeks van fouten, door de andere maanden eraan toe te voegen hetgeen op een herhaling slaat van het hierboven vermelde voorbeeld van het eerste kwartaal. Het is echter duidelijk dat wanneer informatieverzamelingen geordend zijn men kan vooropstellen dat $\sigma_N^2 < \sigma_D^2 < \sigma_J^2$ namelijk dat een kleinere variantie mag verwacht worden voor die maanden waar meer informatie beschikbaar is. Gegeven de hypothese omtrent informatieverzamelingen bekomen we geen witte ruis proces vanuit theoretisch standpunt. De hypothese omtrent de informatieverzamelingen is echter helemaal niet bindend. Het is even goed plausibel te stellen dat $I_N = I_D = I_J$, namelijk DULBEA beschikt over een massa gegevens die de eerste publikaties genereren. Onjuistheid over

gegevens omtrent januari beïnvloeden nietemin het bepalen van het cijfer van november. De hypothese valt te verdedigen en verdient daarom aandacht. Nu hebben we echter dat fouten, wanneer zij gemaakt werden vanuit eenzelfde informatieverzameling, onderling gekorreleerd zijn (1). We hebben dus in dit geval :

$$\epsilon_D = \rho_1 \epsilon_N + \eta$$

$$\epsilon_J = \delta_1 \epsilon_N + \delta_2 \epsilon_D + \eta \quad (2.2)$$

$$\eta \sim N(0, \sigma_\eta^2)$$

Dit patroon herhaalt zich alweer voor de andere kwartalen zodat wij opnieuw een proces bekomen dat indruist tegen de witte ruis hypothese. De konklusie uit het voorgaande is dat onafhankelijk van hoe informatie gestructureerd is, rationaliteit niet impliceert dat fouten witte ruis zijn. Vermits het ook vrij evident is dat witte ruis fouten niet de rationaliteit van DULBEA impliceren is het dus bewezen dat rationaliteit noch voldoende noch nodige voorwaarde is voor witte ruis fouten.

Deze bevinding heeft evenwel een aantal onaangename gevolgen voor wie rationaliteit wenst te testen. Ekonometrische toesten hoeven nu meer inhoud te geven aan de informatiestructuur. Nemen wij bijvoorbeeld aan dat informatieverzamelingen geordend zijn wat een zekere vorm van heteroskedasticiteit tot gevolg heeft zoals hierboven werd aangetoond. Het stelsel (2.1) kan simultaan worden geschat waarbij de gegevens worden gerangschikt per informatieniveau; een concept geïntroduceerd in G-V [2, p. 416] . Het stelsel (2.2) kan per informatieniveau geschat worden, i.e. alle maanden in de zelfde situatie als januari ten opzichte van eerste

(1) Dit is een klassiek resultaat uit de voorspellingstheorie, zie bijvoorbeeld Granger en Newbold [4, p. 121] .

publicatie tijdstip worden samengebracht, we bekomen dan zoals in (2.2).

$$\text{BNP} = \alpha + \beta E(\text{BNP} | I^t) + \delta_1 \epsilon_1 + \delta_2 \epsilon_2 + \eta$$

Hetgeen een lineaire regressie is met MA (2) storingsterm. Een schattingsprocedure geïntroduceerd door Lars Hansen (2) lost dit probleem op. Rationaliteit impliceert dan $\alpha = 0$ en $\beta = 1$. Naast deze eerste theoretische correctie moet er nog een tweede gemaakt worden die wij aan het eind van de volgende paragraaf uiteenzetten.

3. ENKELE HERINTERPRETATIES VAN RESULTATEN

Er werden twee studies door G-V uitgevoerd. De studie die 1964-1973 besloeg was gekenmerkt door fouten structuren die hooft genaamd niet witte ruis waren. In een latere studie die 1974-77 beslaat werden echter wel witte ruis processen geïdentificeerd. Wanneer de auteur en J. Vuchelen echter [3, p. 57] opmerken dat een systematiek in het foutenpatroon wijst op een irrationaliteit van de opsteller van de reeksen dan hadden zij het verkeerd voor. Niettemin wijzen de statistische resultaten uit tabel 3 in [3] op een witte ruis proces. Zoals in een appendix aan laatst vernoemd artikel wordt duidelijk gemaakt is het vrij moeilijk om empirisch witte ruis vast te stellen. De meeste toetsen belichten een bepaald facet van witte ruis. Wij hebben net aangetoond dat, in geval van rationaliteit, we ons aan een zekere vorm van "heteroskedasticiteit" mogen verwachten voor zover deze term hier geschikt is. Het is gemakkelijk te begrijpen dat met de kleine steekproef waarmee G-V in hun tweede artikel werkten men moeilijk dit fenomeen kon waarnemen. De eerste studie, die 1963-1974 bes-

(2) Voor toepassing en referentie van deze techniek zie Brown en Maital [1].

loeg, en dus op een heel wat grotere steekproef is gebaseerd toonde aan dat fouten zeer sterk gestructureerd waren structuren geïdentificeerd met de Box en Jenkins methodiek die hoegenaamd niets gemeen hebben met deze gesuggereerd in dit artikel. Men mag gerust stellen dat DULBEA over deze periode niet erg rationeel is geweest in de berekeningen van de BNP-indexcijfers. Tabel 4 in G-V [2, p. 416] mag teven als indicatie dienen van de bescheven "heteroskedasticiteit" die natuurlijk ook optreedt bij niet rationele voorspellingen.

De voornaamste konklusie die G-V maakten blijft echter onaangeroerd door de huidige bevinding. De grootorde der fouten wordt kleiner vergeleken over de twee studies. Kleinere fouten zijn een vermeldenswaardig feit dat echter, zoals G-V opmerken [3, p. 71], geen ondubbelzinnige interpretatie heeft. Namelijk het is niet duidelijk of kleinere fouten het gevolg zijn van betere berekeningstechnieken of een louter minder aandacht besteden aan definitieve berekeningen. Deze opmerking, heeft ook haar grote implicaties op de interpretatie van het concept "rationaliteit" in deze kontekst. Immers, rationaliteit van voorspellingen slaat op een situatie waarbij de voorspeller zelf geen invloed heeft op de uiteindelijke realisatie van het proces. DULBEA bepaalt zelf het definitieve cijfer zodat extreem genomen er zich een "moral hazard" - probleem stelt dat elke notie van "rationaliteit" principieel ongeldig maakt. Het begrip "rationaliteit" is hier dus eigenlijk totaal ongepast, wij zouden daarom de term efficiëntie van voorlopige publikaties als een beter begrip willen introduceren.

BIBLIOGRAFIE

- [1] BROWN, B.W. and MAITAL, S. (1981) : "What do economists know ? An empirical study of experts' expectations" *Econometrica* Vol. 49 nr 2.
- [2] GHYSELS, E. en VUCHELEN, J. (1980) : "Fouten in de voorlopige BNP-cijfers van DULBEA : een tijdreeksanalyse" *Cahiers Economiques de Bruxelles* nr 87.
- [3] GHYSELS, E. en VUCHELEN, J. (1981) : "Het gebruik van de DULBEA-BNP cijfers in konjunkturanalyses" *Cahiers Economiques de Bruxelles* nr 89.
- [4] GRANGER, C. and NEWBOLD, P. (1977) : "Forecasting Economic Time Series" Academic Press.

SUMMARY

In previous work the author and J. Vuchelen [2] and [3] errors, i.e. the difference between final and preliminary releases, of GNP data published by DULBEA were analyzed. In the present article it is proven that "rationality" a concept related to the expected value of GNP conditional upon all available information, of preliminary publications is neither necessary nor sufficient for errors to behave like white noise as is frequently assumed and wrongly supposed in our previous work. The main conclusion from our previous work though still applies. Finally as a matter of semantics we claim that the term "rationality" is somewhat misplaced in this type of literature because the final figure is determined by the same body determining the preliminary and hence errors are entirely endogeneous. We think that the concept "efficiency" of preliminary releases is a more appropriate concept.