SAMENVATTING (SUMMARY IN DUTCH)

De methode voor autoregressieve-voortschrijdende gemiddelde (vanaf nu ARMA) tijdreeksanalyse, zoals die geïntroduceerd is in het invloedrijke boek van Box en Jenkins in 1970, wordt vaak gebruikt bij het bouwen van dynamische economische modellen. Deze aanpak omvat drie stappen: de identificatie, het schatten, en de diagnostische beoordeling van een model. De Box-Jenkins methode van modelkeuze in tijdreeksen voldoet echter niet altijd. Zo is bijvoorbeeld uit simulatiestudies gebleken dat het juiste model regelmatig niet wordt teruggevonden, en ook wordt er te vaak verondersteld dat een tijdreeks een eenheidsswortel bevat. Men kan zich voorstellen dat dit grote gevolgen kan hebben voor verdere stappen in het dynamische modelbouwproces.

In dit proefschrift behandel ik onderwerpen die te maken hebben met de modelkeuze voor tijdreeksen met en zonder seizoensgedrag. De aandacht gaat vooral uit naar geschikte modelkeuzestrategieën, die alternatief zijn aan de traditionele Box-Jenkins aanpak. Tal van praktijkvoorbeelden en Monte Carlo simulaties dienen ter illustratie.

Hoofdstuk 2 is gewijd aan een kort overzicht van de tijdreeksanalyse in het tijdsdomein, waarbij ik de concepten en notatie hanteer zoals die in het boek van Box en Jenkins zijn gegeven. De begrippen integratie en cointegratie worden iets gedetailleerder beschreven. Hierbij staat de methode van Johansen (1988) centraal.

Hoofdstuk 3 behandelt modelkeuze voor univariate tijdreeksen zonder seizoensgedrag. De Box-Jenkins en gerelateerde methoden worden kritisch geëvalueerd. Uit een paar Monte Carlo simulaties blijkt bijvoorbeeld dat de Hannan en Rissanen (1982) methode vaak niet in staat is het juiste model terug te vinden. Andere simulaties laten zien dat de keuze tussen twee incorrecte modellen wordt beïnvloed door het onbekende data genererende proces. De frequentie van verwerping van deze modellen via een toets op de aanwezigheid van residuele autocorrelatie blijkt echter niet door dit ware model te worden bepaald. Dit leidt tot de formulering van een alternatieve modelkeuzestrategie, met als stappen: de specificatie en het schatten van een verzameling van modellen, het diagnostisch toetsen van al deze modellen op eventuele gebreken, en het kiezen van een model uit de resterende verzameling modellen. Deze selectie is gebaseerd op de waarden van verscheidene criteria, welke gerelateerd zijn aan het doel van de modelbouw. Deze fasen

van modelkeuze worden geïllustreerd aan de hand van de tijdreeks van het werkloosheidspercentage in de Verenigde Staten, een voorbeeld dat als een rode draad door hoofdstuk 3 loopt.

Een belangrijke, maar vaak lastige, beslissing die genomen moet worden heeft betrekking op de stationariteit van een tijdreeksvariabele. Als geconstateerd wordt dat een reeks niet stationair is, dan wordt deze vaak getransformeerd door eerste verschillen te nemen. Een nadeel hiervan kan zijn dat een reeks wordt overgedifferentiëerd, hetgeen betekent dat een eenheidswortel wordt geïntroduceerd aan de MA kant van een ARMA model. Deze wortel zorgt ervoor dat parameterschattingen afwijkingen vertonen en dus onbruikbaar zijn om op eenheidswortels te toetsen. In paragraaf 3.3 stel ik een toetsmethode voor die niet aan afwijkingen onderhevig is. Deze methode is gebaseerd op de autocorrelaties van de getransformeerde tijdreeks en levert standaard normaal verdeelde toetsgrootheden op. Een toepassing op enkele reeksen uit Schwert (1987) leidt tot de suggestie dat twee keer differentiëren van een tijdreeks, zoals hij dat doet, soms niet nodig is.

Om de eigenschap van stationariteit van een univariate tijdreeks te onderzoeken beperkt men zich in de regel tot univariate tijdreeksmodellen. Gegeven de vele mogelijke onderlinge relaties tussen economische variabelen is het denkbaar dat een univariate reeks deel uit maakt van een systeem van variabelen dat middels een multivariaat tijdreeksmodel beschreven kan worden. De maximale orde van een ARMA model voor een univariate reeks kan worden afgeleid van de orde van het multivariate model. Toetsgrootheden in de univariate context zijn echter vaak gebaseerd op AR benaderingen. In paragraaf 3.3 laat ik analytisch zien dat inferentie op basis van een te klein gekozen AR model - een vergissing die gauw gemaakt is - kan leiden tot de onjuiste suggestie van de aanwezigheid van eenheidswortels. Dit wordt empirisch geïllustreerd met de bekende Lydia Pinkham dataset.

Nadat men een besluit heeft genomen over de stationariteit van een reeks, kan men overgaan tot de constructie van een verzameling modellen. Een inspectie van grafieken en de, eventueel recursieve, autocorrelaties komen bij deze constructie van pas. De modellen worden allen getoetst op de achterliggende veronderstellingen, zoals bijvoorbeeld witte ruis storingen, normaliteit, en de afwezigheid van afwijkende waarnemingen.

Een formele behandeling van modelkeuzecriteria vindt plaats in de laatste paragraaf van hoofdstuk 3. Deze criteria worden verdeeld in twee categorieën, die gerelateerd zijn aan het doel waarvoor het model gebruikt wordt, i.e. voorspellen en beschrijven. Ik stel een nieuw voorspelcriterium

voor en pas deze tevens toe. Soms bepalen enkele waarnemingen de keuze van een model. Een methode voor de opsporing van dit soort waarnemingen wordt hier kort beschreven en eveneens toegepast. Tot besluit van hoofdstuk 3, neem ik de keuze tussen een eerste orde AR en een eerste orde MA model onder de loep. Een nieuw keuzecriterium, dat gebaseerd is op de geschatte meervoudige correlatiecoëfficiënt, blijkt beter zijn doel te bereiken dan een elders voorgesteld criterium.

Hoofdstuk 4 is grotendeels gewijd aan univariate seizoenstijdreeksen. De meeste aandacht gaat uit naar modellen voor de seizoenspatronen in niet-stationaire reeksen. Voorbeelden van deze modellen zijn die waarbij uit gegaan wordt van de transformatie naar eerste verschillen (met dummy variabelen voor de seizoensschommelingen), of naar jaarlijkse verschillen. Box en Jenkins stelden voor om zelfs beide transformaties te hanteren. De keuze tussen deze transformaties is lastig omdat de meeste huidige toets-grootheden uitgaan van verschillende nulhypothesen, die ook nog eens niet-stationair zijn. Dat laatste heeft tot gevolg dat tabellen met kritieke waarden voor de toetsgrootheden moeten worden gegenereerd via Monte Carlo simulaties. In paragraaf 4.1 stel ik een alternatieve modelkeuzestrategie voor die uitgaat van een stationair model, die tal van modellen als speciaal geval heeft. De idee is om de autocorrelaties van het zogeheten 'airline model' te onderzoeken op indicaties van overdifferentiatie. Dit model, dat voorgesteld is door Box en Jenkins, blijkt in de praktijk voor veel tijdreeksen te voldoen. De toetsgrootheden blijken nu asymptotisch standaard normaal verdeeld te zijn.

De bruikbaarheid van deze toetsprocedure is echter beperkt: het hangt af van de geschiktheid van het airline model. Een methode die complexere modellen toelaat is ontwikkeld door Hylleberg en anderen (1990) [HEGY]. Het jaarlijkse verschillenfilter veronderstelt de aanwezigheid van zogeheten seizoenseenheidswortels. De HEGY methode is toegepast op kwartaaldata in het artikel van Hylleberg en anderen, en uitbreidingen hiervan naar (twee-) maandelijkse cijfers zijn beschreven in paragraaf 4.2. Tabellen met kritieke waarden van de toetsgrootheden staan vermeld in appendix B.

Het belang van het toetsen op eenheidswortels wordt onderstreept in paragraaf 4.3. De aandacht gaat uit naar de keuze tussen een, zo te noemen, FDSD model waarin eerste orde verschillen worden gecombineerd met seizoens-dummies, en het airline model van Box en Jenkins (vanaf hier een MSBJ model genoemd). Het blijkt dat op basis van de conventionele toetsen op residuele autocorrelaties het moeilijk is de beide modellen te onderscheiden. Onder

andere via simulaties laat ik zien dat het hanteren van een MSBJ model, terwijl een FDSD model adequaat is, systematische fouten bij het voorspellen tot gevolg heeft. Verder blijkt dat de constructie van een multivariaat tijdreeksmodel ook wordt vergemakkelijkt als voor iedere betrokken reeks een geschikt model voor de seizoenspatronen wordt gebruikt. Tevens probeer ik te onderbouwen dat het beschouwen van een MSBJ-achtig model kan leiden tot de onjuiste suggestie van de aanwezigheid van niet-lineaire patronen. Dit kan worden veroorzaakt door een aantal additieve afwijkende waarnemingen in een tijdreeks, die door een FDSD model kan worden beschreven.

Een geheel nieuwe methode voor modelselectie in, mogelijkerwijs niet-stationaire, seizoenstijdreeksen wordt voorgesteld in de laatste paragraaf van hoofdstuk 4. De idee is om, bijvoorbeeld, een kwartaalreeks weer te geven middels een multivariaat autoregressief model voor de vector die de vier jaarreeksen met de waarnemingen per kwartaal bevat. Het blijkt dat de aan- of afwezigheid van, eventueel gerestricteerde, cointegratierelaties een indicatie geeft voor een geschikt univariaat model. Het op deze wijze naar een seizoensreeks kijken introduceert tevens een nieuwe vorm van grafische weergave. De figuren van de vier jaarreeksen per kwartaal kunnen alvast een vermoeden geven welk model wellicht wel of niet in aanmerking zal komen. In het tweede deel van deze paragraaf wordt deze multivariate aanpak gegeneraliseerd naar bivariate tijdreeksen.

Hoofdstuk 5 sluit dit proefschrift af met de conclusies en een kort overzicht van de belangrijkste resultaten.

The Tinbergen Institute is the Netherlands Research Institute and Graduate School for General and Business Economics founded by the Faculties of Economics (and Econometrics) of the Erasmus University in Rotterdam, the University of Amsterdam and the Free University in Amsterdam. The Tinbergen Institute, named after the Nobel prize laureate professor Jan Tinbergen, has the responsibility for the Ph.D.-program of the three faculties mentioned.

Copies of the books which are published in the Tinbergen Institute Series can be ordered through Thesis Publishers, Bickersgracht 60, 1013 LG Amsterdam, The Netherlands, phone: + 3120 6255429.

The following books already appeared in this series:

Subseries A. General Economics

no. 1 Otto H. Swank, "Policy Makers, Voters and Optimal Control. Estimation of the Preferences behind Monetary and Fiscal Policy in the United States".

no. 2 Jan van der Borg, "Tourism and Urban Development. The Impact of Tourism on Urban Development: Towards a Theory of Urban Tourism, and its Application to the Case of Venice, Italy".

no. 3 Albert Jolink, "Liberté, Egalité, Rareté. The Evoluationary Economics of Léon Walras".

no. 5 Rudi M. Verburg, "The Two Faces of Interest. The Problem of Order and the Origins of Political Economy and Sociology as Distinctive Fields of Inquiry in the Scottish Enlightenment".

no. 6 Harry P. van Dalen, "Economic Policy in a Demographically Divided World".

no. 8 Marian Hofkes, "Modelling and Computation of General Equilibrium".

no. 12 Kwame Nimako, "Economic Change and Political Conflict in Ghana 1600-1900".

no. 13 Ans Vollering, "Care Services for the Elderly in The Netherlands".

no. 15 Cees Gorter, "The Dynamics of Unemployment and Vacancies on Regional Labour Markets".

no. 16 Paul Kofman, "Managing Primary Commodity Trade (On the Use of Futures Markets)".

no. 18 Philip Hans Franses, "Model Selection and Seasonality in Time Series".

Subseries B. Business Economics

no. 4 Rob Buitendijk, "Towards an Effective Use of Relational Database".

no. 7 P.J. Verbeek, "Two Case Studies on Manpower Planning in an Airline".

no. 9 T.C.R. van Someren, "De Rol van de Organisatorische Vernieuwingen".

no. 10 M. van Vliet, "Optimization of Manufacturing System Design".

no. 11 R.M.C. van Waes, "Architectures for Information Management. A Pragmatic Approach on Architectural Concepts and their Application in Dynamic Environments".

no. 14 Shuzhong Zhang, "Stochastic Queue Location Problems".

no. 17 Paul Th. van de Laar, "Financieringsgedrag in de Rotterdamse Maritieme Sector, 1945-1960".