VU Research Portal

Armax model specification testing, with an application to unemployment in the Netherlands

Bierens, H.J.

1986

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Bierens, H. J. (1986). *Armax model specification testing, with an application to unemployment in the Netherlands*. (Serie Research Memoranda; No. 1986-26). Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

05348

1986

SERIE RESEARCH MEMORANDA

ARMAX MODEL SPECIFICATION TESTING, WITH AN
APPLICATION TO UNEMPLOYMENT IN THE NETHERLANDS

By Herman J. Bierens

Researchmemorandum 1986-26

Oktober 1986

VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
A M S T E R D A M

ARMAX MODEL SPECIFICATION TESTING, WITH AN APPLICATION TO UNEMPLOYMENT IN THE NETHERLANDS

by Herman J. Bierens^{*)}

VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN

Postbus 7161 - 1007 MC Amsterdam

Abstract: In this paper we present a new model specification test for ARMAX models. The null hypothesis that the ARMAX model represents the conditional expectation of the dependent variable given the entire past of the economic time series process under review is tested against the alternative hypothesis that the null is false. The test involved is a consistent standard normal test, as under the null hypothesis the test statistic is asymptotically $N(0,1)$ distributed, whereas under the alternative that the null is false and under the hypothesis of strict stationarity of the data generating process the absolute value of the test statistic converges in probability to infinity. This test is applied to testing the rational expectations-natural rate (RE-NR) hypothesis for the Netherlands according to the approach of Sargent (1976), on the basis of monthly data and the ARMAX framework. As for the result, it appears that unemployment in the Netherlands is caused, in the sense of Granger (1969), by the money interest rate and unemployment itself. According to Sargent the RE-NR hypothesis therefore has to be rejected. This empirical result can be explained by Baumol's (1959) sales revenue maximization hypothesis, augmented with the hypothesis of a flexible labor effort rate.

^{*)} This paper has been presented at the 1986 CAMP Econometrics conference and at econometrics seminars at the University of California, San Diego, the University of Minnesota, Queen's University, the University of Western Ontario and the University of Toronto. In particular I acknowledge the comments of Robert Engle, Joop Hartog, Tammo Oegema, Teo Perez, Benedict M. Pötscher, Christopher A. Sims and two referees.

April 1986

1. INTRODUCTION

In this paper we present a new model specification test for testing an ARMAX model specification against the alternative hypothesis that the ARMAX specification involved is false. This test is a further elaboration of the model specification tests of Bierens (1982b, 1984) and Bierens and Hartog (1984)

Under the null hypothesis the test statistic of our test is standard normally distributed. Under the maintained hypothesis of strict stationarity the test is consistent in the sense that the absolute value of the test statistic converges in probability to infinity if the null hypothesis is false. This implies that asymptotically any misspecification will be detected. In the non-stationary case this general alternative hypothesis is no longer manageable, due to the fact that then the true model may no longer be time invariant. Cf. Bierens (1984, Sec. 9). Nevertheless, also in that case the implicit alternative will cover a wide range of misspecifications.

Given a $k+1$ - variate time series process $\{(y_t, x_t)\}$, where y_t and the k components of x_t are real-valued random variables, the ARMAX model assumes the form:

$$(1.1) \quad (1 - \sum_{s=1}^p \alpha_s L^s) y_t = \mu + \sum_{s=1}^q \beta'_s L^s x_t + (1 + \sum_{s=1}^r \gamma_s L^s) e_t,$$

where L is the usual lag operator (thus $L^s y_t = y_{t-s}$, $L^s x_t = x_{t-s}$, etc.), $\mu \in R$, $\alpha_s \in R$, $\beta_s \in R^k$ and $\gamma_s \in R$ are unknown parameters, the e_t 's are the errors and p , q and r are natural numbers specified in advance. The exclusion of x_t in this model is no loss of generality, as we may replace x_t by $x_t^* = x_{t+1}$. Now the null hypothesis to be tested is

$$(1.2) \quad H_0: E(e_t | (y_{t-1}, x_{t-1}), (y_{t-2}, x_{t-2}), \dots) = 0 \quad \text{a.s. for each } t,$$

or in words: for each t the mathematical expectation of the error conditional on the entire past of the process equals zero with probability 1. Assuming that the lag polynomial $1 + \sum_{s=1}^r \gamma_s L^s$ is invertible, this hypothesis implies that the ARMAX model (1.1) represents the mathematical expectation of y_t conditional on the entire past of the process. Thus if (1.2) holds the model represents the best forecasting scheme in terms of a quadratic loss function, as then the forecast error variance is minimal.

The ARMAX model specification is particularly suitable for macroeconomic vector time series modeling without imposing *a priori* restrictions prescribed by macroeconomic theory. Such macroeconomic analysis has been advocated and conducted by Sims (1980, 1981) and Doan, Litterman and Sims (1984) in the framework of unrestricted vector autoregressions and observable index models. Cf. Sargent and Sims (1977) for the latter models. These models are special cases of ARMAX models. The advantage of ARMAX models over the models used by Sims is that ARMAX models allow an infinite lag structure with a parsimonious parametrization, by which we get a tractable model that may better reflect the strong dependence of macroeconomic time series.

Estimation of the parameters of a vector ARMAX model for the case that the x_t 's are exogenous has been considered by Hannan, Duinsmuir and Deistler (1980), among others. This estimation theory, however, is not straightforwardly applicable to the model under review, for the following three reasons. Firstly, the condition that the x_t 's are exogenous and thus that the error process (e_t) is independent of the process (x_t) is too strong a condition, as then feedback from y_t to x_t is excluded. Secondly, incorporating this condition in the test procedure would affect the type I error of the test, as then the test may reject the true null hypothesis (1.2) because of failure of the auxiliary hypothesis involved. For the very same reason we do not assume that the errors e_t are Gaussian or independent. Thirdly, the ARMAX structure need not hold under the alternative hypothesis that (1.2) is false. So we need to embed this model in a more general data generating process in order to keep track of the asymptotic properties of the parameter estimates and the test statistic under the alternative hypothesis. This situation is similar to that in Bierens (1984), where we proposed a model specification test for nonlinear autoregressions, i.e., before we can test the model specification we need first to develop a suitable estimation theory.

Our test is particularly suitable as a first stage in testing the rational expectations- natural rate (RE-NR) hypothesis, along the lines of Sargent's (1976) approach. According to Sargent (1976), for testing the RE-NR-hypothesis it is sufficient to test whether output and unemployment, respectively, are caused, in the sense of Granger (1969),

by no other variables. Thus if y_t in (1.1) is the unemployment rate then the RE-NR-hypothesis just corresponds to the case that all the β_s 's are zero vectors, as then the x_t 's can be deleted from the information set without affecting the forecasting power of the model. The Granger causality concept, however, requires that the model represents the conditional expectation of y_t given the entire past of the process $\{(y_t, x_t)\}$. Therefore one should first test whether the null hypothesis (1.2) holds before Sargent's test can be applied on the basis of the ARMAX specification under review.

The above outlined approach has been used to test the RE-NR-hypothesis for the Netherlands on the basis of monthly data. In the first instance we specified an ARMAX model with y_t the level of unemployment, seasonally adjusted and x_t a vector with components the money interest rate, the growth of industrial production, the growth of export, the growth of the wage rate and the inflation rate. This specification, however, was rejected by the test. After replacing y_t with the log of the unemployment level the test accepted an ARMAX specification of the data generating process. In conducting Sargent's test it appeared that the money interest rate is significantly Granger-causing unemployment. Thus the RE-NR-hypothesis has to be rejected for the Netherlands. As will be argued this empirical result can be explained by Baumol's (1959) sales revenue maximization hypothesis, augmented with the hypothesis of a flexible labor effort rate.

The emphasis of this paper is on the application of the ARMAX model specification test. Technical details of the estimation and test procedures are given in the Appendices. The main text of the paper aims to be accessible for the general reader. Therefore, in Section 2 only an informal outline of the estimation and test procedures is given which is just sufficient for understanding the empirical application in Section 3.

2. AN INFORMAL OUTLINE OF THE ESTIMATION AND TEST PROCEDURES

2.1 The Model

We recall that the ARMAX model (1.1) represents the conditional expectation of y_t given the entire past of the process $\{(y_t, x_t)\}$, provided condition (1.2) holds and the MA lag polynomial $1 + \sum_{s=1}^r \gamma_s L^s$ is invertible. We then have

$$(2.1) \quad E(y_t | (y_{t-1}, x_{t-1}), (y_{t-2}, x_{t-2}), \dots) \\ = \mu / (1 + \sum_{s=1}^r \gamma_s) + \{(\sum_{s=1}^p \alpha_s L^s + \sum_{s=1}^r \gamma_s L^s) / (1 + \sum_{s=1}^r \gamma_s L^s)\} y_t \\ + \{(\sum_{s=1}^q \beta'_s L^s) / (1 + \sum_{s=1}^r \gamma_s L^s)\} x_t \quad \text{a.s.}$$

and

$$(2.2) \quad e_t = y_t - E(y_t | (y_{t-1}, x_{t-1}), (y_{t-2}, x_{t-2}), \dots) \quad \text{a.s.}$$

Since the MA lag polynomial can be written as $1 + \sum_{s=1}^r \gamma_s L^s = \prod_{s=1}^r (1 - \lambda_s L)$, where $\lambda_1^{-1}, \dots, \lambda_r^{-1}$ are its possibly complex-valued roots, invertibility requires $|\lambda_s| < 1$ for $s=1, \dots, r$. In particular, for $0 < \delta < 1$ the set

$$(2.3) \quad \Gamma_\delta = \{(\gamma_1, \dots, \gamma_r)' \in R^r : |\lambda_s| \leq 1-\delta \text{ for } s=1, \dots, r\}$$

is a compact set of vectors $(\gamma_1, \dots, \gamma_r)'$ with this property. The compactness of Γ_δ follows from the fact that the γ_s 's are continuous functions of $\lambda_1, \dots, \lambda_r$, hence Γ_δ is the continuous image of a compact set and therefore compact itself. Cf. Royden (1968, Proposition 4, p. 158).

From now on we assume that there are known compact subsets M of R , A of R^p , B of $R^{k \cdot q}$ and Γ_δ of R^r such that, if (1.2) is true,

$$(2.4) \quad \mu \in M, (\alpha_1, \dots, \alpha_p)' \in A, (\beta'_1, \dots, \beta'_q)' \in B \text{ and } (\gamma_1, \dots, \gamma_r)' \in \Gamma_\delta$$

Stacking all these parameters in a vector θ_0 :

$$(2.5) \quad \theta_0 = (\mu, \alpha_1, \dots, \alpha_p, \beta'_1, \dots, \beta'_q, \gamma_1, \dots, \gamma_r)'$$

and denoting

$$(2.6) \quad \Theta = M \times A \times B \times \Gamma_{\delta},$$

which is a compact subset of R^m with $m = 1 + p + kq + r$, we thus have $\theta_0 \in \Theta$ if (1.2) is true.

Now denote

$$(2.7) \quad z_t = (y_t, x_t')'.$$

Then the conditional expectation (2.1) can be written as

$$(2.8) \quad E(y_t | z_{t-1}, z_{t-2}, \dots) = \phi(\theta_0) + \sum_{s=1}^{\infty} \eta_s(\theta_0)' z_{t-s},$$

where

$$(2.9) \quad \phi(\theta_0) = \mu / (1 + \sum_{s=1}^r \gamma_s)$$

and the $\eta_s(\cdot)$ are continuously differentiable vector functions defined by:

$$(2.10) \quad (1 + \sum_{s=1}^r \gamma_s L^s) (\sum_{s=1}^{\infty} \eta_s(\theta_0) L^s) = \begin{bmatrix} \sum_{s=1}^p \alpha_s L^s + \sum_{s=1}^r \gamma_s L^s \\ \sum_{s=1}^q \beta_s L^s \end{bmatrix}$$

It is not too hard to verify that each component $\eta_{i,s}(\theta)$ of $\eta_s(\theta)$ satisfies

$$(2.11) \quad \sum_{s=1}^{\infty} \sup_{\theta \in \Theta} |\eta_{i,s}(\theta)| < \infty,$$

$$(2.12) \quad \sum_{s=1}^{\infty} \sup_{\theta \in \Theta} |(\partial/\partial \theta_j) \eta_{i,s}(\theta)| < \infty \quad (j=1, 2, \dots, m)$$

$$(2.13) \quad \sum_{s=1}^{\infty} \sup_{\theta \in \Theta} |(\partial/\partial \theta_j)(\partial/\partial \theta_l) \eta_{i,s}(\theta)| < \infty \quad (j, l=1, 2, \dots, m)$$

etc. These properties will play a crucial role in our estimation theory. In particular, the model (2.1) can now be written as a nonlinear regression model:

$$(2.14) \quad y_t = g_t(\theta_0) + e_t,$$

where the response function

$$(2.15) \quad g_t(\theta) = \phi(\theta) + \sum_{s=1}^{\infty} \eta_s(\theta)' z_{t-s}$$

and its first and second partial derivatives are well-defined random functions.

2.2 Estimation

Assuming that only $z_1, \dots, z_n$ have been observed, we now propose to estimate θ_0 by nonlinear least squares, as follows. Let

$$(2.16) \quad \begin{aligned} \hat{g}_t(\theta) &= \phi(\theta) + \sum_{s=1}^{t-1} \eta_s(\theta)' z_{t-s} && \text{if } t \geq 2, \\ &= \phi(\theta) && \text{if } t \leq 1 \end{aligned}$$

and denote

$$(2.17) \quad \hat{Q}(\theta) = (1/n) \sum_{t=1}^n \{y_t - \hat{g}_t(\theta)\}^2.$$

Then the proposed least squares estimator $\hat{\theta}$ of θ_0 is a (measurable) solution of

$$(2.18) \quad \hat{\theta} \in \Theta : \hat{Q}(\hat{\theta}) = \inf_{\theta \in \Theta} \hat{Q}(\theta).$$

Similar to the results in Bierens (1984) we can set forth conditions such that under the null hypothesis (1.2),

$$(2.19) \quad \sqrt{n} (\hat{\theta} - \theta_0) \rightarrow N_m(0, \Omega_1^{-1} \Omega_2 \Omega_1^{-1}) \text{ in distr.,}$$

where Ω_1 is the probability limit of

$$(2.20) \quad \hat{\Omega}_1 = (1/n) \sum_{t=1}^n \{(\partial/\partial \theta') \hat{g}_t(\hat{\theta})\} \{(\partial/\partial \theta) \hat{g}_t(\hat{\theta})\}$$

and Ω_2 is the probability limit of

$$(2.21) \quad \hat{\Omega}_2 = (1/n) \sum_{t=1}^n \{y_t - \hat{g}_t(\hat{\theta})\}^2 \{(\partial/\partial\theta') \hat{g}_t(\hat{\theta})\} \{(\partial/\partial\theta) \hat{g}_t(\hat{\theta})\}.$$

Moreover, when the null hypothesis (1.2) is false we show that there exists a $\theta_* \in \Theta$ such that

$$(2.22) \quad \text{plim}_{n \rightarrow \infty} \hat{\theta} = \theta_*.$$

The conditions and proofs of these results are reminiscent of those in Bierens (1984) and will be given in Appendix B.

2.3 The model specification test

In Bierens (1984) we have proposed a consistent model specification test for nonlinear time series regressions. This test tests the null hypothesis that the errors of the nonlinear regression model under review obey a condition of the type (1.2). Therefore this model specification test is in principle also applicable to the ARMAX case considered in this paper. A disadvantage of this test, however, is that the distribution of the test statistic under the null hypothesis is of an unknown type, so that the critical region of the test involved had to be derived on the basis of Chebishev's inequality for first absolute moments. This approach will lead, of course, to overestimating the effective type I error of the test, as Chebishev's inequality is not very sharp. Moreover, this test is quite laborious for relatively large data sets and models. On the other hand, the test involved is consistent in the sense that any model misspecification will be detected as the sample size grows to infinity, provided the data generating process is strictly stationary. To the best of our knowledge no other model specification test for time series regressions has this consistency property.

We shall now propose a new test which has a known limiting distribution under the null hypothesis and is consistent in the above sense. In particular, we shall construct a test statistic $\sqrt{n}\hat{T}$, say, with the property that under the null hypothesis (1.2), $\sqrt{n}\hat{T} \rightarrow N(0,1)$ in distribution as $n \rightarrow \infty$, whereas under the alternative hypothesis that (1.2) does

not hold and under the stationarity hypothesis, $\text{plim}_{n \rightarrow \infty} |\sqrt{n}\hat{T}| = \infty$.

As in Bierens (1984), this new test is based on the identification of the null hypothesis in terms of unconditional moments, as follows.

THEOREM 1. Let $\{(v_t, z_t)\}$ be a strictly stationary stochastic process in $R \times R^m$ with the following properties:

(I) Let $E|v_t| < \infty$; $E\|z_t\|^1 < \infty$.

(II) There exists a strictly stationary α -mixing process (u_t) in a Euclidean space U and a Borel measurable mapping Λ from the space of one-sided infinite sequences in U into $R \times R^m$, such that

$$(2.23) \quad (v_t, z_t')' = \Lambda(u_t, u_{t-1}, u_{t-2}, \dots) \quad \text{a.s.}$$

Let $z_{i,t}$ be the i -th component of z_t , let $z_{i,t}^{(\ell)}$ be the value of $z_{i,t}$ rounded off to ℓ decimal digits and let

$$(2.24) \quad z_t^{(\ell)} = (z_{1,t}^{(\ell)}, \dots, z_{m,t}^{(\ell)})'.$$

Finally, let ψ be an arbitrary bounded continuous one-to-one mapping from R into R . There exist a subset S of $R \times R^m \times (-1, 1)$ with Lebesgue measure zero and a natural number ℓ_0 such that

$$(2.25) \quad P\{E(v_t | z_{t-1}, z_{t-2}, \dots) = 0\} < 1$$

if and only if for each $(\rho, \xi, \tau) \in R \times R^m \times (-1, 1) - S$ and each natural number $\ell \geq \ell_0$,

$$(2.26) \quad E v_t \cdot \exp \left\{ \rho \psi \left(\sum_{j=1}^{\infty} \tau^{j-1} \xi' z_{t-j}^{(\ell)} \right) \right\} \neq 0.$$

Proof: Appendix C.

For the reader who is not familiar with the α -mixing and Borel measurability concepts we note that the α -mixing concept allows the process to have an asymptotically vanishing memory, and Borel measurability allows the function involved to be highly discontinuous, as long as

this function yields a properly defined random variable if we substitute random variables for its arguments. For a discussion of the α -mixing concept, see Appendix A and, e.g., White and Domowitz (1984). For Borel measurability, see e.g. Chung (1974) and Royden (1968). The zero Lebesgue measure of the set S implies that if we draw ρ , τ and the components of ξ randomly from say the uniform $(-1,1)$ distribution, then

$$(2.27) \quad P((\rho, \xi, \tau) \in S) = 0,$$

hence Theorem 1 implies that for some natural number ℓ_0 and all $\ell \geq \ell_0$,

$$(2.28) \quad P\{E(v_t | z_{t-1}, z_{t-2}, \dots) = 0\} < 1$$

if and only if

$$(2.29) \quad P\{E\{v_t \exp\{(\sum_{j=1}^{\infty} \tau^{j-1} \xi' z_{t-j}^{(\ell)})\} | \rho, \xi, \tau\} = 0\} < 1.$$

Now let us introduce some notation:

$$(2.30) \quad v_t = y_t - g_t(\text{plim}_{n \rightarrow \infty} \hat{\theta}); \quad (\text{cf. (2.15), (2.18), (2.19), (2.22)})$$

$$(2.31) \quad \hat{v}_t = y_t - \hat{g}_t(\hat{\theta}) \quad ; \quad (\text{cf. (2.16), (2.18)})$$

$$(2.32) \quad z_t = (y_t, x_t')';$$

$$(2.33) \quad d_{t,\ell}(\rho, \xi, \tau) = \exp\{\rho \psi(\sum_{j=1}^{\infty} \tau^{j-1} \xi' z_{t-j}^{(\ell)})\}; \quad (\text{cf. (2.26)})$$

$$(2.34) \quad \begin{aligned} \hat{d}_{t,\ell}(\rho, \xi, \tau) &= \exp\{\rho \psi(\sum_{j=1}^{t-1} \tau^{j-1} \xi' z_{t-j}^{(\ell)})\} & \text{if } t \geq 2, \\ &= 1 & \text{if } t \leq 1; \end{aligned}$$

$$(2.35) \quad c_{n,\ell}(\rho, \xi, \tau) = (1/n) \sum_{t=1}^n v_t d_{t,\ell}(\rho, \xi, \tau);$$

$$(2.36) \quad \hat{c}_{n,\ell}(\rho, \xi, \tau) = (1/n) \sum_{t=1}^n \hat{v}_t \hat{d}_{t,\ell}(\rho, \xi, \tau),$$

where $z_t^{(\ell)}$ is defined as in Theorem 1. Theorem 1 says that the null hypothesis (1.2) is false if and only for each $\ell \geq \ell_0$ and each $(\rho, \xi, \tau) \in \mathbb{R}^m \times (-1,1) - S$,

$$(2.37) \quad \bar{c}_\ell(\rho, \xi, \tau) = \lim_{n \rightarrow \infty} E c_{n, \ell}(\rho, \xi, \tau) \neq 0.$$

It can be shown that $\hat{c}_{n, \ell}(\rho, \xi, \tau)$ is a consistent estimator of $\bar{c}_\ell(\rho, \xi, \tau)$. The conclusion (2.37) therefore suggests to use $\hat{c}_{n, \ell}(\rho, \xi, \tau)$ as a basis for a consistent test of the hypothesis (1.2). In particular, it can be shown that under the null hypothesis (1.2),

$$(2.38) \quad \text{plim}_{n \rightarrow \infty} \{ \sqrt{n} \hat{c}_{n, \ell}(\rho, \xi, \tau) - (1/\sqrt{n}) \sum_{t=1}^n v_t(\hat{d}_{t, \ell}(\rho, \xi, \tau) - \bar{b}_\ell(\rho, \xi, \tau)' \Omega_1^{-1}(\partial/\partial \theta') \tilde{g}_t(\text{plim}_{n \rightarrow \infty} \hat{\theta})) \} = 0,$$

where Ω_1 is the probability limit of $\hat{\Omega}_1$ defined by (2.20) and

$$(2.39) \quad \bar{b}_\ell(\rho, \xi, \tau) = \lim_{n \rightarrow \infty} (1/n) \sum_{t=1}^n E d_{t, \ell}(\rho, \xi, \tau) (\partial/\partial \theta') g_t(\text{plim}_{n \rightarrow \infty} \hat{\theta}).$$

Denoting

$$(2.40) \quad \bar{s}_\ell^2(\rho, \xi, \tau) = \lim_{n \rightarrow \infty} (1/n) \sum_{t=1}^n E v_t^2 \{ d_{t, \ell}(\rho, \xi, \tau) - \bar{b}_\ell(\rho, \xi, \tau)' \Omega_1^{-1}(\partial/\partial \theta') g_t(\text{plim}_{n \rightarrow \infty} \hat{\theta}) \}^2.$$

it follows now from (2.38) and the martingale central limit theorem of Brown (1971) (as quoted by Bierens (1984)) that under the null hypothesis (1.2)

$$(2.41) \quad \sqrt{n} \hat{c}_{n, \ell}(\rho, \xi, \tau) \rightarrow N(0, \bar{s}_\ell^2(\rho, \xi, \tau)) \text{ in distribution}$$

for each $(\rho, \xi, \tau) \in R \times R^m \times (-1, 1)$ and each natural number ℓ .

It will also be shown that

$$(2.42) \quad \hat{s}_{n, \ell}^2(\rho, \xi, \tau) = (1/n) \sum_{t=1}^n \hat{v}_t^2 \{ \hat{d}_{t, \ell}(\rho, \xi, \tau) - \hat{b}_{n, \ell}(\rho, \xi, \tau)' \hat{\Omega}_1^{-1}(\partial/\partial \theta') \tilde{g}_t(\hat{\theta}) \}^2$$

with

$$(2.43) \quad \hat{b}_{n,l}(\rho, \xi, \tau) = (1/n) \sum_{t=1}^n \hat{d}_{t,l}(\rho, \xi, \tau) (\partial/\partial \theta') \hat{g}_t(\hat{\theta})$$

is a consistent estimator of $\bar{s}_l^2(\rho, \xi, \tau)$, regardless whether or not the null is true. Hence, denoting

$$(2.44) \quad \hat{T}_l(\rho, \xi, \tau) = \hat{c}_{n,l}(\rho, \xi, \tau) / \sqrt{(\hat{s}_{n,l}^2(\rho, \xi, \tau))}$$

we have under the null hypothesis

$$(2.45) \quad \sqrt{n} \hat{T}_l(\rho, \xi, \tau) \rightarrow N(0,1) \text{ in distr.}$$

for each $(\rho, \xi, \tau) \in R \times R^m \times (-1,1)$ and each natural number l . Moreover, if the null is false then (2.37) implies

$$(2.46) \quad \text{plim}_{n \rightarrow \infty} |\sqrt{n} \hat{T}_l(\rho, \xi, \tau)| = \infty$$

for each $(\rho, \xi, \tau) \in R \times R^m \times (-1,1) - S$ and each $l \geq l_0$.

In practice the exceptional set S is unknown. However, as argued before, if we draw $\hat{\rho}$, $\hat{\tau}$ and the components of $\hat{\xi}$ randomly from say the uniform $(-1,1)$ distribution and if we substitute $(\hat{\rho}, \hat{\xi}, \hat{\tau})$ for (ρ, ξ, τ) then the result (2.45) goes through under the null and (2.46) goes through under the alternative that (1.2) is false, provided that in the latter case l is large enough.

In the practical application in Section 3 our test is applied in a slightly modified version, for the following reason. Suppose we have chosen for the continuous bounded function ψ , (cf.(2.26))

$$(2.47) \quad \psi(.) = \text{atan}(\delta(.)), \text{ where } \delta > 0 \text{ is some constant.}$$

Moreover, suppose that

$$(2.48) \quad \hat{w}_{t,l}(\tau, \xi) = \sum_{j=1}^{t-1} \tau^{j-1} \xi' z_{t-j}^{(l)}$$

takes with high probability only large positive values. Then $\psi(\hat{w}_{t,l}(\tau, \xi))$ takes with high probability only values close to $\frac{1}{2}\pi$, which is the upperbound of the function ψ . But then the function

$$(2.49) \quad \hat{d}_{t,l}(\rho, \xi, \tau) = \exp(\rho \psi(\delta \hat{w}_{t,l}(\tau, \xi)))$$

(cf. (2.34)) takes values close to $\exp(\rho \frac{1}{2} \pi)$ and consequently

$$\hat{c}_{n,l}(\rho, \xi, \tau) \sim \{(1/n) \sum_{t=1}^n \hat{v}_t\} \exp(\rho \frac{1}{2} \pi).$$

But since the $\hat{v}_t$ are least squares residuals of a model with a constant term they sum up to zero and hence $\hat{c}_{n,l}(\rho, \xi, \tau)$ will be close to zero. Clearly this will destroy the power of the test. Therefore we propose to standardize the argument of ψ , i.e., we propose to replace $\hat{w}_{t,l}(\tau, \xi)$ in (2.49) by $\hat{\bar{w}}_{t,l}(\tau, \xi)$ defined as follows:

$$(2.50) \quad \hat{\bar{w}}_{t,l}(\tau, \xi) = \{\hat{w}_{t,l}(\tau, \xi) - (1/t) \sum_{s=1}^t \hat{w}_{s,l}(\tau, \xi)\} \\ / \{(1/t) \sum_{s=1}^t \hat{w}_{s,l}(\tau, \xi)^2 - ((1/t) \sum_{s=1}^t \hat{w}_{s,l}(\tau, \xi))^2\}^{\frac{1}{2}} \\ = 0 \quad \begin{array}{l} \text{if } t \geq 3, \\ \text{if } t \leq 2, \end{array}$$

where $\hat{w}_{t,l}(\tau, \xi) = 0$ if $t \leq 1$ and defined by (2.45) if $t \geq 2$. Choosing $\psi(\cdot)$ as in (2.47) the function $\hat{d}_{t,l}(\rho, \xi, \tau)$ is thus redefined as

$$(2.51) \quad \hat{d}_{t,l}(\rho, \xi, \tau) = \exp(\rho \cdot \text{atan}(\delta(\hat{\bar{w}}_{t,l}(\tau, \xi)))) \quad \begin{array}{l} \text{if } t \geq 3, \\ = 1 \quad \text{if } t \leq 2. \end{array}$$

Redefining the function $d_{t,l}(\rho, \xi, \tau)$ accordingly as

$$(2.52) \quad d_{t,l}(\rho, \xi, \tau) = \exp(\rho \cdot \text{atan}(\delta(\bar{w}_{t,l}(\tau, \xi)))) ,$$

where

$$(2.53) \quad \bar{w}_{t,l}(\tau, \xi) = \{w_{t,l}(\tau, \xi) - E w_{t,l}(\tau, \xi)\} / \{\text{var}(w_{t,l}(\tau, \xi))\}^{\frac{1}{2}}$$

with

$$(2.54) \quad w_{t,l}(\tau, \xi) = \sum_{j=1}^{\infty} \tau^{j-1} \xi_{t-j}^{(l)},$$

all the previous results go through. This modified test with $\delta = 0.5$ will be used in the empirical application in Section 3.

The conditions and proofs of these results will be given in Appendix C.

Remark: In Bierens (1985) it is argued that the parameters τ and ξ for which $E(v_t | z_{t-1}^{(l)}, z_{t-2}^{(l)}, \dots) = E(v_t | \sum_{j=1}^{\infty} \tau^{j-1} \xi' z_{t-j}^{(l)})$ a.s. are likely to be irrational. Since this result plays a key-role in the proof of Theorem 1, one might therefore think that the consistency of our test can not hold in practice as consistency requires irrational τ and ξ , whereas in practice it is impossible to deal with irrational numbers. However, the function

$$(2.55) \quad \bar{T}_\ell(\rho, \xi, \tau) = \bar{c}_\ell(\rho, \xi, \tau) / \sqrt{\bar{s}_\ell^2(\rho, \xi, \tau)} = \text{plim}_{n \rightarrow \infty} \hat{T}_\ell(\rho, \xi, \tau)$$

is continuous in (ρ, ξ, τ) , hence if $\bar{T}_\ell(\rho, \xi, \tau) \neq 0$ holds then it holds too on an open neighborhood of (ρ, ξ, τ) , and thus also for all rational (ρ_*, ξ_*, τ_*) in this neighborhood.

2.4 Extensions

Although the test in Section 2.3 has been derived for testing a linear ARMAX model specification, it can easily be extended to testing non-linear ARMAX models of the form

$$(2.56) \quad (1 + \sum_{s=1}^r \gamma_s L^s) y_t = f(y_{t-1}, \dots, y_{t-p}, x_{t-1}, \dots, x_{t-q}, \theta_0^{(1)}) \\ + (1 + \sum_{s=1}^r \gamma_s L^s) e_t,$$

where the null hypothesis is the same as in (1.2). In this model the function $f(\cdot)$ is continuous in its arguments. The vector $\theta_0^{(1)}$ is a vector of unknown parameters (different from (2.5)) in a compact subset $\Theta^{(1)}$ of a Euclidean space. This model encompasses the nonlinear ARX model in Bierens (1984), as the latter model can be obtained from (2.56) by setting $\gamma_s = 0$ for $s = 1, \dots, r$.

Similarly to (2.8) we can write:

$$(2.57) \quad E(y_t | z_{t-1}, z_{t-2}, \dots)$$

$$= (1 + \sum_{s=1}^r \gamma_s L^s)^{-1} f(y_{t-1}, \dots, y_{t-p}, x_{t-1}, \dots, x_{t-q}, \theta_0^{(1)}) = \sum_{s=0}^{\infty} \eta_s^* (z_{t-s}^*, \theta_0),$$

say, provided (1.2) holds, where

$$(2.58) \quad z_t^* = (y_{t-1}, \dots, y_{t-p}, x_{t-1}', \dots, x_{t-q}')',$$

$$(2.59) \quad \theta_0 = (\theta_0^{(1)}, \gamma_1, \dots, \gamma_r) \in \Theta = \Theta^{(1)} \times \Gamma_\delta.$$

Redefining $g_t(\theta)$ and $\tilde{g}_t(\theta)$ as

$$(2.60) \quad g_t(\theta) = \sum_{s=0}^{\infty} \eta_s^* (z_{t-s}^*, \theta),$$

$$(2.61) \quad \tilde{g}_t(\theta) = \begin{cases} \sum_{s=0}^{t-2} \eta_s^* (z_{t-s}^*, \theta) & \text{for } t \geq 2, \\ 0 & \text{for } t \leq 1, \end{cases}$$

we can set forth mild conditions such that the argument in Sections 2.2 and 2.3 goes through.

We shall not give a formal proof of this proposition. We only provide a key-theorem (Theorem A4) in Appendix A which will enable us to modify the argument in Appendices B and C to the nonlinear ARMAX model under review. The modification itself is left to the reader.

3. THE GRANGER-CAUSAL STRUCTURE OF UNEMPLOYMENT IN THE NETHERLANDS

3.1 Testing the RE-NR-hypothesis: a review

In this section we shall apply the approach in Section 2 to monthly Dutch data, where emphasis is placed on unemployment. Our primary aim is to test the RE-NR-hypothesis for the Netherlands.

The rational expectations (RE) hypothesis concerns the idea that expectations of economic agents are essentially conditional expectations, where the conditioning is on all available information about the economic process involved. Thus if $\{(y_t, x_t)\}$ is this economic process then (2.1) represents the rational expectation of y_t , provided the null hypothesis (1.2) holds. This RE-hypothesis is due to Muth (1961), but its rise is due to Lucas (1972, 1973), Sargent (1973, 1976) and Sargent and Wallace (1975), among others, who combined the RE concept with the natural rate (NR) hypothesis of Phelps (1967) and Friedman (1968). The NR-hypothesis states that the deviation of the actual level of output and the unemployment rate from their natural rates is merely due to the difference between the actual inflation rate and the expected inflation rate, where the natural rates involved are exogenous variables. Combining the NR-hypothesis with the RE-hypothesis yields a far reaching consequence for the effectivity of economic policy, namely the consequence that deterministic monetary and fiscal policy rules have no lasting impact on output and unemployment. In other words: monetary and fiscal policy rules are impotent.

Because of its radical policy implication it is very important to test the RE-NR-hypothesis empirically. Direct tests of the RE-hypothesis have been conducted by Pesando (1975), Carlson (1977), Mullineaux (1978) and Figlewski and Wachtel (1981), using Livingston's data on price expectations. The results of these tests do not give conclusive support to the RE-hypothesis. Tests of the combined RE-NR-hypothesis or its implied impotence hypothesis have been conducted by Lucas (1972, 1973), Sargent (1973, 1976), Barro (1977, 1978, 1979), Barro and Rush (1980) and Pesaran (1982). Except Pesaran, these authors claimed to have found empirical evidence in favor of the RE-NR-hypothesis. Their tests involve testing restrictions implied by the RE-NR-hypothesis on the coefficients of the models they are working with. However, the RE-hypothesis requires that these models represent conditional expectation functions, where the

conditioning is on the whole past of the economic process under review, but none of these authors have tested this crucial hypothesis properly. As a consequence, finding the coefficient of a variable violating the RE-NR-hypothesis to be insignificant does not provide enough evidence in support of the hypothesis involved. Pesaran has tested Barro's version of the RE-NR-hypothesis against some non-nested Keynesian alternatives, using Barro's data and applying the non-nested test procedure proposed by Pesaran and Deaton (1978). His conclusion is that neither groups of hypotheses can reject the other. This result, however, indicates that either the test involved has not enough power or that it is likely that all the models are misspecified, where the misspecification may concern the hypotheses of interest or some auxiliary hypotheses such as the linearity hypothesis.

Our test of the RE-NR-hypothesis is based on Sargent's definition of the natural rate of unemployment hypothesis. Quoting Sargent (1976), "the natural rate hypothesis does not rule out the possibility that there are correlations between the unemployment rate and other variables such as prices or wages or money supply. It does imply, however, that any such correlations that exist cannot be exploited in predicting the unemployment rate". Thus the RE-NR-hypothesis implies "that the unemployment rate is caused, in Granger's, sense, by no other variables". (Recall that according to Granger (1969), we say that y_t is causing x_t if we are better able to predict x_t using all available past information, including past y_t , than if the information apart from past y_t had been used.) The test we shall conduct is similar to Sargent's test, except that we use the ARMAX framework (1.1) with $k=5$ and $p=q=r=1$, where y_t is measure of unemployment in the Netherlands and the 5 components of x_t are other variables relevant for the performance of the Dutch economy. Thus we employ the most parsimonious ARMAX model, i.e.,

$$(3.1) \quad y_t = \mu + \alpha y_{t-1} + \sum_{i=1}^5 \beta_i x_{i,t-1} + e_t + \gamma e_{t-1}.$$

We recall that the RE-NR-hypothesis corresponds to the hypothesis

$$(3.2) \quad H_0: \beta_i = 0 \quad \text{for } i = 1, \dots, 5,$$

provided condition (1.2) holds. So our test is in fact a two-stage test: first we test whether the ARMAX model is acceptable, using the test proposed in Section 2.3, and then we test by means of the well-known Wald test whether the hypothesis (3.2) is acceptable.

3.2 The data

Our data set (Source: CBS (1960-1983)) covers the period 1960-1 through 1982-12 and consists of monthly observations on the following six variables:

- u = level of unemployment, in thousands of labor years, seasonally adjusted.
- r = mortgage interest rate (percentage)
- p = price index of finished products
- o = volume of industrial output
- e = volume of exports
- w = gross hourly wages.

The main reason for working with monthly data concerns the length of the series involved. The estimation and test procedures in Section 2 are only asymptotically valid; thus the more observations we have, the better the asymptotic approximations will be. On the other hand, however, we have to face the problem that only a limited number of variables are available on a monthly basis. So the analysis we are going to conduct is only a partial analysis and therefore we have to be careful in drawing too general economic conclusions from our results.

Although the approach in Section 2 allows for a fair amount of heterogeneity of the data generating process, time trends are usually not allowed. Time trends converging to infinity as time tends to infinity are likely to occur in the series with respect to the variables p, o, e and w but not in u and r. The unemployment level is typically a bounded variable, bounded from below by zero and from above by the total labor force. Thus the series involved cannot contain a time trend converging to infinity. The same applies to the interest rate r: we can hardly imagine interest rates growing to infinity. On the other hand price and wage inflation and the growth of production and export levels

seem to be of persisting nature, which indicates exponential time trends in the levels of these series. In order to remove time trends and possible seasonal patterns in these four variables we have transformed them into annual percentage changes. Thus denoting the levels of p , o , e or w in month t by x_t , the transformation involved is

$$(3.3) \quad \dot{x}_t = 100 (x_t - x_{t-12}) / x_{t-12}.$$

In the sequel we shall denote this transformation by a dot. Thus the transformed data set now consists of 264 monthly observations on the variables $y_t = u_t$ and $x_t = (x_{1,t}, \dots, x_{5,t})' = (r_t, \dot{o}_t, \dot{e}_t, \dot{p}_t, \dot{w}_t)'$.

3.3 Estimation and test results

Our first attempt to estimate the unemployment function according to the approach in Section 2.2 yields the following result.

$$(3.4) \quad u = -0.1671 + 1.015 u_{-1} - 0.00403 r_{-1} + 0.3627 \dot{p}_{-1} - 0.09405 \dot{o}_{-1} \\ (-0.049) \quad (149.7) \quad (-0.0082) \quad (1.733) \quad (-0.9663) \\ -0.03321 \dot{e}_{-1} - 0.06556 \dot{w}_{-1} \\ (-0.7338) \quad (-0.396) \\ \gamma = 0.4281 ; R^2 = .9975 ; SE = 6.553 ; n=263 \\ (5.436)$$

These results include the coefficient of determination R^2 , the residual standard error SE and the number of observations n used for estimation. The numbers between brackets are the t -values. The γ is the MA parameter.

If one would accept this model specification as being correct, at first sight the conclusion would be that the RE-NR hypothesis is confirmed, as the Wald test of the hypothesis that all the variables except u_{-1} can be deleted yields a χ^2_5 statistic with value 5.99. However, the coefficient of u_{-1} is larger than 1, hence the model is unstable and therefore implausible. This conclusion is confirmed by the model specification test. The ARMAX model specification test has been conducted in tenth-fold, where in each run the test parameters $(\hat{\rho}, \hat{\xi}, \hat{\tau})$ have been drawn independently from the uniform distribution on the hypercube $(-1,1) \times (X_{i=1}^{k+1}(0.1,1)) \times (0,0.9)$. Moreover, the rounding off parameter ℓ

has been chosen equal to 1. We also conducted this series of tests for $l = 2, 3, 4, 5$, but the test results involved were practically the same. The test results are the following:

ARMAX model specification test results for u

-4.213 -3.976 -3.271 -3.795 -3.214 -3.653 -0.165 -3.935 -1.209 -2.635

These test results indicate that the ARMAX model specification involved is suspect. With two exceptions the test statistics are outside any reasonable confidence interval of the standard normal distribution. Thus no conclusions can be drawn from the above estimation results.

It should be noted that the test runs do not yield independent test statistics. So it is not quite correct a procedure to count the number of test statistics outside say the 95% confidence interval of the standard normal distribution in order to determine whether the results are due to a type I error or not. The proper way of conducting the test is to draw $(\hat{\rho}, \hat{\xi}, \hat{\tau})$ only once and merely decide on the basis of the resulting test statistic. On the other hand, however, we feel that conducting the test several times may better reveal possible model misspecification.

The question now arises how to repair this misspecification. One way would be to choose a richer parametrisation and then to try again. In the present case, however, intuition suggests a different approach. The intuitive idea is that the rejection of the above model is due to too large a variation in the unemployment series in comparison with the variation in the other five time series, so that dampening the unemployment series by taking a log-transformation might have a favorable effect. Indeed, replacing u by $\ln(u)$ gives a much better result.

$$\begin{aligned}
 (3.5) \quad \ln(u) = & 0.0496 + 0.9244 \ln(u_{-1}) + 0.03974 r_{-1} - 0.00017 \dot{p}_{-1} \\
 & (0.642) (49.12) \quad (3.480) \quad (-0.069) \\
 & -0.00128 \dot{o}_{-1} - 0.00018 \dot{e}_{-1} - 0.00040 \dot{w}_{-1} \\
 & (-0.6004) (-0.1719) (0.1798) \\
 \gamma = & 0.07598 ; R^2 = 0.9836 ; SE = 0.1303 ; n = 263. \\
 & (0.6916)
 \end{aligned}$$

ARMAX model specification test results for $\ln(u)$

0.5856 0.6943 1.0980 0.7853 0.5153 0.0002 0.6722 0.3367 -0.4007 0.2935

The specification test, which is conducted in the same way as before, clearly accepts the ARMAX specification at any reasonable significance level.

Since the estimate of the MA parameter γ is insignificant, the model reduces to an ARX model. So we have next set $\gamma = 0$ and re-estimated the model. Also, we repeated the model specification test in the same way as before. It is not surprising that also the ARX model appears to be acceptable. The estimation and test results are the following.

$$(3.6) \ln(u) = 0.0430 + 0.9322 \ln(u_{-1}) + 0.03601 r_{-1} - 0.00029 \dot{p}_{-1} \\ (0.5864) (56.64) (3.416) (-0.1215) \\ -0.00132 \dot{o}_{-1} + 0.00002 \dot{e}_{-1} - 0.00025 \dot{w}_{-1} \\ (-0.6273) (0.0140) (-0.1207)$$

$$R^2 = 0.9832 ; SE = 0.1316 ; n = 263.$$

ARX model specification test results for $\ln(u)$

0.3986 0.5968 0.9215 0.6524 0.4439 -0.1925 0.5972 0.1778 -0.5152 -0.1765

We see that the parameter of r_{-1} is strongly significant. This violates the RE-NR hypothesis, as the money interest rate is Granger-causing the log of unemployment. A formal test of the RE-NR hypothesis is the Wald test of the hypothesis (3.2). This test yields a test statistic with value 23.91, which is far outside any reasonable confidence interval of the χ^2_5 distribution. According to Sargent (1976) the RE-NR hypothesis should therefore be rejected for the Netherlands.

Testing the null hypothesis that all variables except $\ln(u_{-1})$ and r_{-1} can be deleted from the model yields a χ^2_4 test statistic with value 0.6284, and can therefore not be rejected at any reasonable significance

level. Deleting these variables now yields the following simple regression model:

$$(3.7) \ln(u) = 0.0253 + 0.9300 \ln(u_{-1}) + 0.0383 r_{-1}$$

(0.561) (57.24) (4.453)

$$R^2 = 0.9825 \quad ; \quad SE = 0.1353; \quad n = 275.$$

3.4 Interpretation of the results

Our finding that the interest rate is significantly Granger-causing unemployment does not correspond to any established macro-economic theory. At first sight one might think of an effect of interest on unemployment via a Keynesian investment schedule. If this interpretation is correct, however, one also should find a significant impact of the growth of industrial production on unemployment, as investment goods are to a large extent supplied by the industrial sector. Since the growth of industrial production is not significantly Granger-causing unemployment, this interpretation has to be rejected. Also, the neoclassical theory of the firm does not provide an explanation of our finding, because a change in the interest rate affects the fixed cost of firms and not their variable cost (except the cost of holding inventories), hence marginal cost will not or hardly change. Since in the neoclassical theory of the firm the marginal cost curve is just the supply curve of the firm, interest will have no substantial effect on output and prices (except if the interest rate affects the marginal revenue curve). Baumol's (1959) sales revenue maximisation theory, however, augmented with the concept of a flexible labor effort rate defined below, do provide a possible explanation.

Baumol's hypothesis is that firms are sales revenue maximisers rather than profit maximisers. Profit only plays a role as a constraint, i.e. sales revenue is maximised subject to the constraint that profit does not fall below some lower bound. Thus if $R(X)$ denotes the total revenue function and $C(X)$ denotes the total cost function, then the firm chooses its output level X_0 as to

$$(3.8) \quad R(X_0) = \max R(X),$$

subject to

$$(3.9) \quad R(X) - C(X) \geq \bar{\pi},$$

where $\bar{\pi}$ is the profit constraint. Now suppose that the firm's short term production function is of the Leontief type:

$$(3.10) \quad X = \min\{\alpha \ell L, \beta \bar{K}\},$$

where L is the firm's demand for labor, $\bar{K}$ is the capital stock, ℓ is the labor effort rate, and α and β are positive valued parameters. The labor effort rate ℓ expresses the simple fact that if workers work harder they can produce more, and vice versa. It is assumed that the labor effort rate cannot exceed an upperbound $\bar{\ell}$: if for given L , $X/(\alpha L)$ increases beyond $\bar{\ell}$ then employment will increase from L to $\bar{L} = X/(\alpha \bar{\ell})$. Moreover, we assume that the firm will only lay off workers if profit falls below the profit constraint $\bar{\pi}$. This assumption is consistent with Baumol's hypothesis, for if firms are not profit maximisers then neither are they cost minimisers. Furthermore, this assumption also corresponds to the often observed fact that firms do not adjust their employment smoothly, but only reorganise production and lay off workers when they are forced to by losses.

Now let w be the wage rate, let r be the interest rate and let p_k be the price index of capital goods at the time the firm bought its capital stock $\bar{K}$. Moreover, let L_0 be the initial labor force, let

$$(3.11) \quad R(X^*) = \max R(X)$$

and assume

$$(3.12) \quad X^* \leq \beta \bar{K}, \quad X^*/(\alpha L_0) \leq \bar{\ell},$$

$$(3.13) \quad R(X^*) - wL_0 - rp_k \bar{K} \geq \bar{\pi}.$$

Furthermore, assume for the moment that the sales revenue function $R(X)$ does not depend on w and r . We shall now consider the effect of an increase of r and w , respectively.

Consider first the case that r increases to r' . It is clear that output and employment do not change as long as the profit constraint remains non-binding. Now suppose

$$(3.14) \quad R(X^*) - wL_0 - r'p_k\bar{K} < \bar{\pi}.$$

If

$$(3.15) \quad R(X^*) - wX^*/(\alpha\bar{l}) - r'p_k\bar{K} \geq \bar{\pi}$$

then the firm is able to restore its production on the sales revenue maximising output level X^* with profit not below the lower bound $\bar{\pi}$ by reducing its labor force from L_0 to L' , where

$$(3.16) \quad X^*/(\alpha\bar{l}) \leq L' \leq (R(X^*) - r'p_k\bar{K} - \bar{\pi})/w.$$

In that case output is not affected by the interest rate increase, but employment is reduced.

If the increase in r is so strong that

$$(3.17) \quad R(X^*) - wX^*/(\alpha\bar{l}) - r'p_k\bar{K} < \bar{\pi}$$

then output will decrease to the level X' where

$$(3.18) \quad R(X') - wX'/(\alpha\bar{l}) - r'p_k\bar{K} = \bar{\pi},$$

provided such an output level X' is feasible. In that case employment will decrease to

$$(3.19) \quad L' = X'/(\alpha\bar{l})$$

Finally, if

$$(3.20) \quad \max_X \{R(X) - wX/(\alpha\bar{l}) - r'p_k\bar{K}\} < \bar{\pi}$$

and if $\bar{\pi}$ is just the normal profit the firm has to earn in order to survive, the firm will quit business.

Note that if we make the assumption that in the case (3.16) the firm will choose the lowest level of L' , the demand for labor of the firm becomes:

$$\begin{aligned}
 (3.21) \quad L &= L_0 && \text{if } R(X^*) - wL_0 - rp_k \bar{K} \geq \bar{\pi}, \\
 &= X^*/(\alpha \bar{L}) && \text{if } R(X^*) - wL_0 - rp_k \bar{K} < \bar{\pi} \\
 &&& \text{and } R(X^*) - wX^*/(\alpha \bar{L}) - rp_k \bar{K} \geq \bar{\pi}, \\
 &= X'/(\alpha \bar{L}) && \text{if } R(X^*) - wX^*/(\alpha \bar{L}) - rp_k \bar{K} < \bar{\pi} \\
 &&& \text{and } R(X') - wX'/(\alpha \bar{L}) - rp_k \bar{K} = \bar{\pi}, \\
 &= 0 && \text{if } \max_X \{R(X) - wX/(\alpha \bar{L}) - rp_k \bar{K}\} < \bar{\pi}.
 \end{aligned}$$

The case that w increases to w' is similar to above. By changing r' to r and w to w' in (3.14) through (3.20) the same conclusions can be drawn. However, the impact of w on employment is not found back, as the percentage change of the real wage rate is not significantly Granger causing unemployment. A possible explanation for this is that the sales revenue function is not independent of w , but usually shifts upward with w . If the wage rate increase is due to the outcome of a collective bargaining process (as applies to about 75% of all Dutch wage earners) income and expenditure of most wage earners will increase too, which causes an upward shift of the revenue functions of firms. This upward shift may offset the upward shift of cost functions and hence neutralise the effect of a wage rate increase on employment.

Also the interest rate may have an income effect on sales revenue, via interest earnings from savings minus interest payments on debts. Moreover, a high interest rate will have a negative impact on the demand for durables. It seems likely that the negative income and expenditure effects are dominant for low income earners, whereas the positive income effect will be dominant for high income earners. The net effect of the interest rate on sales revenue will therefore depend on the income distribution. Anyhow, the observed positive effect of interest on unemployment suggests that the effect of interest on revenue does not offset the effect of interest on cost.

The theory also explains why we have not found the growth of export volume and industrial production significantly Granger causing unemployment. Consider for example an autonomous upward shift of demand. The sales revenue curve $R(X)$ will then shift upward too and the revenue maximising output level X^* will increase. However, as long as the labor effort rate remains below its maximum $\bar{l}$ employment will remain constant, as firms first exhaust the labor effort rate before they hire new personnel. A fall in demand will shift the revenue function and the revenue maximising output X^* downward. This will only result in layoffs if the profit constraint becomes binding. Thus the output-employment tradeoff seems loose and may therefore vanish at the aggregate level.

Finally, it should be stressed that the above argument does not claim to provide a thorough economic theory of unemployment. Economic reality is much too complex for that. Moreover, we only have focussed on one equation of a system of equations without taking possible cross-equation feedbacks into account. So our analysis is merely a partial analysis serving a limited aim, namely to test the rational expectations-natural rate hypothesis for the Netherlands. Rational expectations theorists might possibly argue that our test is of no use at all, as the observed relationship between unemployment and the money interest rate is just a nonsense relation due to a common trend in the time series involved. The argument in this subsection, however, does intend to refute that point of view.

APPENDIX A: UNIFORM WEAK LAWS

In this appendix we state and prove a number of theorems which we need for proving the results in this paper. The latter proofs will be given in Appendices B and C or their extended versions, which are available from the author on request. We start with the definitions of the α and ϕ -mixing concepts, the stochastic stability concept, the v -stability concept, the concept of proper convergence of distribution functions, and the proper heterogeneity concept.

Let (u_t) be a sequence of random variables in a Euclidean space U . Let $F_{-\infty}^t$ be the σ -field generated by $u_t, u_{t-1}, u_{t-2}, \dots$ and let F_t^∞ be the σ -field generated by $u_t, u_{t+1}, u_{t+2}, \dots$. For $m=0,1,2,\dots$, denote

$$\phi(m) = \sup_t \sup_{E_1 \in F_t^\infty, E_2 \in F_{-\infty}^{t-m}, P(E_2) > 0} |P(E_1 | E_2) - P(E_1)|,$$

$$\alpha(m) = \sup_t \sup_{E_1 \in F_t^\infty, E_2 \in F_{-\infty}^{t-m}} |P(E_1 \text{ and } E_2) - P(E_1)P(E_2)|.$$

DEFINITION A1. The process (u_t) is said to be ϕ -mixing if $\lim_{m \rightarrow \infty} \phi(m) = 0$ and it is said to be α -mixing if $\lim_{m \rightarrow \infty} \alpha(m) = 0$.

For a further discussion of these concepts, see White and Domowitz (1984).

Next, let X be a Euclidean space and let (x_t) be a sequence of X -valued random variables with the structure

$$(A1) \quad x_t = \psi_t(u_t, u_{t-1}, u_{t-2}, \dots),$$

where (u_t) is a sequence of random variables in a Euclidean space U and the ψ_t are Borel measurable mappings from the space of one-sided infinite U -valued sequences into X .

DEFINITION A2. If for $n = 1,2,\dots$ there exist a sequence (m_n) of positive integers and a double array $(x_{t,n})$ of X -valued random variables with the structure

$$x_{t,n} = \psi_{t,n}(u_t, u_{t-1}, \dots, u_{t-m_n+1}),$$

such that together, $m_n = o(n)$ as $n \rightarrow \infty$ and

$$\lim_{n \rightarrow \infty} (1/n) \sum_{t=1}^n (|x_t - x_{t,n}| > \delta) = 0 \quad \text{for every } \delta > 0,$$

then (x_t) is said to be stochastically stable with respect to the base (u_t) .

DEFINITION A3. Let for some $p > 0$, $E||x_t||^p < \infty$, and for $m = 0, 1, 2, \dots$ denote $v(m) = \sup_t E||x_t - E(x_t | u_t, u_{t-1}, \dots, u_{t-m+1})||^p$. If $\lim_{m \rightarrow \infty} v(m) = 0$ then the process (x_t) is said to be v -stable in L^p with respect to the base (u_t) .

For a discussion of the stochastic stability concept, see Bierens (1981, Ch5, 1982a, 1984), and for a discussion of the v -stability concept, see Bierens (1983). Note that v -stability implies stochastic stability. However, the reverse is not generally true.

If (x_t) and the base (u_t) are strictly stationary, the function ψ_t defined by (A1) is time invariant.: $\psi_t = \psi$ for all t . In that case v -stability of x_t in L^p with respect to (u_t) only requires the condition $E||x_t||^p < \infty$.

THEOREM A1. Let (u_t) be a strictly stationary stochastic process in a Euclidean space U , let ψ be a Borel measurable mapping from the space of one-sided infinite sequences in U into a Euclidean space X , and let

$$(A2) \quad x_t = \psi(u_t, u_{t-1}, u_{t-2}, \dots) \quad \text{a.s.}$$

If

$$(A3) \quad E||x_t||^p < \infty$$

then (x_t) is v -stable in L^p with respect to the base (u_t) .

Proof. Let M be a positive constant. From Chung (1974 Theorem 9.4.8) and (A2) it follows that

$$\begin{aligned} (A4) \quad & \lim_{m \rightarrow \infty} E\{x_t I(|x_t| \leq M) | u_t, u_{t-1}, \dots, u_{t-m+1}\} \\ &= E\{x_t I(|x_t| \leq M) | u_t, u_{t-1}, u_{t-2}, \dots\} \\ &= x_t I(|x_t| \leq M) \quad \text{a.s.,} \end{aligned}$$

where $I(\cdot)$ is the indicator function. Since a.s. convergence implies convergence in probability and convergence in probability of bounded random variables implies convergence in L^p (cf. Chung (1974, Theorem 4.5.4)), it follows from (A4) that for every $p > 0$,

$$(A5) \quad \lim_{m \rightarrow \infty} E\{|x_t I(|x_t| \leq M) - E\{x_t I(|x_t| \leq M) | u_t, u_{t-1}, \dots, u_{t-m+1}\}|^p} = 0.$$

Moreover, by Jensen's inequality for conditions expectations we have

$$(A6) \quad E\{|E\{x_t I(|x_t| > M) | u_t, u_{t-1}, \dots, u_{t-m+1}\}|^p} \leq E\{|x_t|^p I(|x_t| > M)\}.$$

Using the easy inequality $|a+b+c|^p \leq 3^p(|a|^p + |b|^p + |c|^p)$ we now conclude from (A5) and (A6)

$$\begin{aligned} (A7) \quad & \limsup_{m \rightarrow \infty} E\{|x_t - E\{x_t | u_t, u_{t-1}, \dots, u_{t-m+1}\}|^p} \\ & \leq 2 \times 3^p E\{|x_t|^p I(|x_t| > M)\}. \end{aligned}$$

for every $M > 0$. Since by (A3) the right hand side of (A7) can be made arbitrarily small by increasing M , (A7) implies

$$\lim_{m \rightarrow \infty} E\{|x_t - E\{x_t | u_t, u_{t-m}, \dots, u_{t-m+1}\}|^p} = 0.$$

This proves the theorem.

Q.E.D.

The next two definitions and corresponding lemmas are needed in order to cope with heterogenous time series processes.

DEFINITION A4. Let (F_n) be a sequence of distribution functions on a Euclidean space X .

- (I) The sequence (F_n) is said to converge properly pointwise in X if there exists a distribution function F on X such that $\lim_{n \rightarrow \infty} F_n = F$ pointwise in the continuity points of F .
- (II) The sequence (F_n) is said to converge properly setwise in X if there exists a distribution function F on X such that for each Borel set B in X , $\lim_{n \rightarrow \infty} \int I(x \in B) dF_n(x) = \int I(x \in B) dF(x)$, where $I(\cdot)$ is the indicator function.

According to Feller (1966, §. 243) the pointwise proper convergence of F_n implies that $\lim_{n \rightarrow \infty} \int \psi(x) dF_n(x) = \int \psi(x) dF(x)$ for every bounded continuous function ψ on X , and vice versa. Moreover, by an easy extension of Theorem 2.2.15 of Bierens (1981) it can be shown:

LEMMA A1. If $F_n \rightarrow F$ properly pointwise in X , then for every continuous function ψ on X satisfying $\sup_n \int |\psi(x)|^{1+\delta} dF_n(x) < \infty$ for some $\delta > 0$, we have: $\lim_{n \rightarrow \infty} \int \psi(x) dF_n(x) = \int \psi(x) dF(x)$.

A more general result holds in the setwise proper convergence case, namely:

LEMMA A2. If $F_n \rightarrow F$ properly setwise in X , then for every Borel measurable real function ψ on X satisfying $\sup_n \int |\psi(x)|^{1+\delta} dF_n(x) < \infty$ for some $\delta > 0$ we have: $\lim_{n \rightarrow \infty} \int \psi(x) dF_n(x) = \int \psi(x) dF(x)$.

Proof: Royden (1968, Proposition 18, p. 232).

Remark: Strictly speaking the use of the Stieltjes integrals in Lemma A2 requires that the function ψ is a pointwise limit of step functions. If this is not the case then the integrals $\int \psi(x) dF_n(x)$ and $\int \psi(x) dF(x)$ should be interpreted as $\int \psi(x) d\mu_n(x)$ and $\int \psi(x) d\mu(x)$, respectively, where μ_n and μ are the probability measures induced by F_n and F .

The next definition extends the proper convergence concept to distributions of one-sided infinite sequences of random variables.

DEFINITION A6. Let (x_t) be a sequence of random variables in a Euclidean space X , and let $F_{t,m}$ be the distribution function of $(x_t, \dots, x_{t-m})$. The process (x_t) is said to be pointwise (setwise) properly heterogenous if there exists a one-sided infinite sequence $(y_t)_{t=0}^{\infty}$ of X -valued random variables such that for $m = 0, 1, 2, \dots$, $(1/n) \sum_{t=1}^n F_{t,m} \rightarrow H_m$ properly pointwise (setwise), where H_m is the distribution function of $(y_0, y_1, \dots, y_m)$. The sequence $(y_t)_{t=0}^{\infty}$ will be called the mean process.

The next theorem extends and improves the lemmas 2 and 3 in Bierens (1984).

THEOREM A2. Let X be a Euclidean space and let (x_t) be an X -valued stochastically stable process with respect to an α -mixing base. Moreover, let Θ be a compact subset of a Euclidean space and let $g(x, \theta)$ be for each $x \in X$ a continuous real function on Θ and for each $\theta \in \Theta$ a Borel measurable real function on X .

(I) If (x_t) is strictly stationary and if $E \sup_{\theta \in \Theta} |g(x_t, \theta)| < \infty$, then

$$\text{plim}_{n \rightarrow \infty} \sup_{\theta \in \Theta} \left| (1/n) \sum_{t=1}^n \{g(x_t, \theta) - E g(x_t, \theta)\} \right| = 0.$$

(II) Let F_t be the distribution function of x_t and assume either $(1/n) \sum_{t=1}^n F_t \rightarrow F$ properly setwise or $(1/n) \sum_{t=1}^n F_t \rightarrow F$ properly pointwise and $g(x, \theta)$ is continuous in both arguments. If for some $\delta > 0$,

$$\sup_n \int (1/n) \sum_{t=1}^n E \sup_{\theta \in \Theta} |g(x_t, \theta)|^{1+\delta} < \infty,$$

then

$$\text{plim}_{n \rightarrow \infty} \sup_{\theta \in \Theta} \left| (1/n) \sum_{t=1}^n g(x_t, \theta) - \int_X g(x, \theta) dF(x) \right| = 0.$$

(III) Moreover, the integral $\int_X g(x, \theta) dF(x)$ is a continuous real function on Θ .

Proof. The proof of this theorem for the ϕ -mixing case can be found in the extended version of Bierens (1984), which is available from the author on request. As pointed out to me by Benedikt M. Pötscher, the reference in the proof of Lemma 3 in Bierens (1984) to Royden (1968, proposition 18, p. 232) (cf. Lemma A2) is not correct, as only pointwise proper convergence was assumed. So Lemma 3 in Bierens (1984) requires either the additional condition that the function is continuous

in both arguments or that the mean of the distribution functions converges properly setwise.

As far as the ϕ -mixing condition is concerned the lemma in Bierens (1982a) plays a key-role. Referring to Lemma 5.4 of Dvoretzky (1972) or Lemma 2.1 of McLeish (1975) instead of the lemma in Bierens (1982a), however, the parts I en II of the theorem under review easily follow. Part III follows from the dominated convergence theorem. Q.E.D.

The following theorem specializes Theorem A2. This theorem, together with the martingale central limit theorem in Bierens (1984), is the basis for our asymptotic results.

THEOREM A3. Let X be a Euclidean space and let (x_t) be an X -valued stochastic process which is stochastically stable with respect to an α -mixing base and pointwise properly heterogenous with mean process $(y_t)_{t=-\infty}^0$. Let $(\gamma_{j,i}(\theta))_{j=0}^{\infty}$, $i=1, \dots, p$, be sequences of continuous mappings from a compact subset Θ of a Euclidean space into X , such that for $i=1, 2, \dots, p$,

$$(A8) \quad \sum_{j=0}^{\infty} \sup_{\theta \in \Theta} \|\gamma_{j,i}(\theta)\| < \infty.$$

Let ψ be a differentiable real function on $\mathbb{R}^p$ such that

$$(A9) \quad \sup_{\|z\| < a} \|(\partial/\partial z)\psi(z)\| = O(a^\mu) \text{ for } a \rightarrow \infty,$$

where $\mu > 0$ is such that for some $\delta > 0$, $\sup_t E\|x_t\|^{1+\mu+\delta} < \infty$. Denote

$$\Gamma_j(\theta) = (\gamma_{j,1}(\theta), \dots, \gamma_{j,p}(\theta)).$$

Then

$$\text{plim}_{n \rightarrow \infty} \sup_{\theta \in \Theta} |(1/n) \sum_{t=1}^n \psi(\sum_{j=0}^{t-1} \Gamma_j(\theta)' x_{t-j}) - E\psi(\sum_{j=0}^{\infty} \Gamma_j(\theta)' y_{-j})| = 0.$$

Moreover, the limit function $E\psi(\sum_{j=0}^{\infty} \Gamma_j(\theta)' y_{-j})$ is a continuous real function on Θ .

Proof: Denote

$$(A10) \quad \rho_j = \max_{i=1, \dots, p} \sup_{\theta \in \Theta} \|\gamma_{j,i}(\theta)\|.$$

Then condition (A8) implies

$$(A11) \quad \sum_{j=0}^{\infty} \rho_j < \infty.$$

Moreover, denote for nonnegative integers m ,

$$(A12) \quad z_t^{(m)}(\theta) = \sum_{j=0}^m \Gamma_j(\theta)' x_{t-j}.$$

We shall now prove Theorem A3 in four steps, each stated in a lemma.

LEMMA A3. There exists a constant K such that for every $n \geq 1$ and every $m \geq 0$, $E \sup_{\theta \in \Theta} \left| (1/n) \sum_{t=1}^n \psi(z_t^{(\infty)}(\theta)) - (1/n) \sum_{t=1}^n \psi(z_t^{(m)}(\theta)) \right| \leq K \sum_{j=m+1}^{\infty} \rho_j$.

LEMMA A4. For every fixed $m \geq 0$,

$$\lim_{n \rightarrow \infty} \sup_{\theta \in \Theta} \left| (1/n) \sum_{t=1}^n \psi\left(\sum_{j=0}^m \Gamma_j(\theta)' x_{t-j}\right) - E\psi\left(\sum_{j=0}^m \Gamma_j(\theta)' y_{-j}\right) \right| = 0.$$

LEMMA A5. There exists a constant K such that for every $m \geq 0$,

$$\sup_{\theta \in \Theta} \left| E\psi\left(\sum_{j=0}^m \Gamma_j(\theta)' y_{-j}\right) - E\psi\left(\sum_{j=0}^{\infty} \Gamma_j(\theta)' y_{-j}\right) \right| \leq K \sum_{j=m+1}^{\infty} \rho_j.$$

LEMMA A6. The function $E\psi\left(\sum_{j=0}^{\infty} \Gamma_j(\theta)' y_{-j}\right)$ is continuous on Θ .

Realizing that (A11) implies $\lim_{m \rightarrow \infty} \sum_{j=m+1}^{\infty} \rho_j = 0$, the theorem under review now easily follows from these four lemmas.

Proof of Lemma A3

Observe from (A10) and (A12) that for $\theta \in \Theta$,

$$\begin{aligned} \left\| z_t^{(\infty)}(\theta) - z_t^{(m)}(\theta) \right\| &\leq \sum_{j=m+1}^{\infty} \left\| \Gamma_j(\theta)' x_{t-j} \right\| \\ &\leq \sum_{j=m+1}^{\infty} \max_{i=1,2,\dots,p} |\gamma_{j,i}(\theta)' x_{t-j}| \leq \sum_{j=m+1}^{\infty} \rho_j \|x_{t-j}\|. \end{aligned}$$

Moreover, by the mean value theorem there exists a $\lambda_{t,m}(\theta) \in [0,1]$ such that

$$\begin{aligned}
 & |\psi(z_t^{(\infty)}(\theta)) - \psi(z_t^{(m)}(\theta))| \\
 &= |(z_t^{(\infty)}(\theta) - z_t^{(m)}(\theta))' (\partial/\partial z) \psi(\lambda_{t,m}(\theta) z_t^{(\infty)}(\theta) + (1-\lambda_{t,m}(\theta)) z_t^{(m)}(\theta))| \\
 &\leq \|z_t^{(\infty)}(\theta) - z_t^{(m)}(\theta)\| \sup_{\|z\|} \|(\partial/\partial z) \psi(z)\| \leq \max(\|z_t^{(\infty)}(\theta)\|, \|z_t^{(m)}(\theta)\|) \|(\partial/\partial z) \psi(z)\| \\
 &\leq \sum_{j=m+1}^{\infty} \rho_j \|x_{t-j}\| \sup_{\|z\|} \|(\partial/\partial z) \psi(z)\| \leq \sum_{j=0}^{\infty} \rho_j \|x_{t-j}\| \|(\partial/\partial z) \psi(z)\|,
 \end{aligned}$$

where the last inequality follows from (A10). According to condition (A9) there exists a constant C such that

$$(A14) \quad \sup_{\|z\| \leq a} \|(\partial/\partial z) \psi(z)\| = Ca^\mu,$$

hence by (A13) and Hölder's and Liapounov's inequalities

$$\begin{aligned}
 E \sup_{\theta \in \Theta} |\psi(z_t^{(\infty)}(\theta)) - \psi(z_t^{(m)}(\theta))| &\leq \\
 &\leq C \cdot E\{(\sum_{j=m+1}^{\infty} \rho_j \|x_{t-j}\|) (\sum_{j=0}^{\infty} \rho_j \|x_{t-j}\|)^\mu\} \\
 &\leq C \{E(\sum_{j=m+1}^{\infty} \rho_j \|x_{t-j}\|)^{1+\mu}\}^{1/(1+\mu)} \{E(\sum_{j=0}^{\infty} \rho_j \|x_{t-j}\|)^{1+\mu}\}^{\mu/(1+\mu)} \\
 &\leq C \{(\sum_{j=m+1}^{\infty} \rho_j)^\mu (\sum_{j=m+1}^{\infty} \rho_j E\|x_{t-j}\|^{1+\mu})\}^{1/(1+\mu)} \\
 &\quad \times \{(\sum_{j=0}^{\infty} \rho_j)^\mu (\sum_{j=0}^{\infty} \rho_j E\|x_{t-j}\|^{1+\mu})\}^{\mu/(1+\mu)} \\
 &\leq C (\sum_{j=0}^{\infty} \rho_j)^\mu \sup_t E\|x_t\|^{1+\mu} (\sum_{j=m+1}^{\infty} \rho_j) = K \sum_{j=m+1}^{\infty} \rho_j,
 \end{aligned}$$

say. This proves the lemma.

Q.E.D.

Proof of Lemma A4. Lemma A4 follows straightforwardly from Theorem A2.

Proof of Lemma A5. From Lemma A1 and the conditions of Theorem A3 it follows

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n E\|z_{t-m}\|^{1+\mu} = E\|y_{-m}\|^{1+\mu},$$

hence

$$\sup_j \geq 0 E ||y_{-j}||^{1+\mu} < \infty.$$

The lemma now follows similarly to Lemma A3.

Q.E.D.

Proof of Lemma A6. Let $\theta_1 \in \Theta$, $\theta_2 \in \Theta$. Similarly to (A13) (with (A14)) we have

$$\begin{aligned} (A16) \quad & |\psi(\sum_{j=0}^{\infty} \Gamma_j(\theta_1)' y_{-j}) - \psi(\sum_{j=0}^{\infty} \Gamma_j(\theta_2)' y_{-j})| \\ & \leq C ||\sum_{j=0}^{\infty} (\Gamma_j(\theta_1) - \Gamma_j(\theta_2))' y_{-j}|| (\sum_{j=0}^{\infty} \rho_j ||y_{-j}||)^{\mu} \\ & \leq C (\sum_{j=0}^{\infty} (\sum_{i=1}^p ||\gamma_{j,i}(\theta_1) - \gamma_{j,i}(\theta_2)||) ||y_{-j}||) (\sum_{j=0}^{\infty} \rho_j ||y_{-j}||)^{\mu}. \end{aligned}$$

Thus similarly to (A15) it follows from (A16),

$$\begin{aligned} & |E \psi(\sum_{j=0}^{\infty} \Gamma_j(\theta_1)' y_{-j}) - E \psi(\sum_{j=0}^{\infty} \Gamma_j(\theta_2)' y_{-j})| \\ & \leq C. (\sum_{j=0}^{\infty} \rho_j)^{\mu} \sup_j \geq 0 E ||y_{-j}||^{1+\mu} \sum_{j=0}^{\infty} \sum_{i=1}^p ||\gamma_{j,i}(\theta_1) - \gamma_{j,i}(\theta_2)||. \end{aligned}$$

Since the $\gamma_{j,i}(\theta)$'s are continuous on Θ , this result proves the lemma.

Q.E.D.

Finally, in the same way it can be shown that the following generalisation of Theorem A3 holds.

THEOREM A4. Let X , Θ , (x_t) , (y_t) and ψ be as in Theorem A3. Let the functions $\gamma_{i,j}(\theta, x)$, $j=0,1,2,\dots$, $i=1,2,\dots,p$, be continuous real functions on $\Theta \times X$, such that

$$\max_{i=1,\dots,p} \sup_{\theta \in \Theta} |\gamma_{j,i}(\theta, x)| \leq \rho_j \bar{b}(x),$$

where

$$\sum_{j=0}^{\infty} \rho_j < \infty$$

and $\bar{b}(x)$ is a continuous real function on X such that

$$\sup_t E \bar{b}(x_t)^{1+\mu+\delta} < \infty.$$

Finally, let

$$\Gamma_j(\theta, x) = (\gamma_{j,1}(\theta, x), \dots, \gamma_{j,p}(\theta, x))'.$$

Then

$$\text{plim}_{n \rightarrow \infty} \sup_{\theta \in \Theta} \left| (1/n) \sum_{t=1}^n \psi \left(\sum_{j=0}^{t-1} \Gamma_j(\theta, x_{t-j}) \right) - E \psi \left(\sum_{j=0}^{\infty} \Gamma_j(\theta, y_{-j}) \right) \right| = 0.$$

Moreover, the limit function $E \psi \left(\sum_{j=0}^{\infty} \Gamma_j(\theta, y_{-j}) \right)$ is continuous on Θ .

APPENDIX B: CONSISTENCY AND ASYMPTOTIC NORMALITY

In this appendix we set forth conditions such that the results in Section 2.2 hold. These conditions are even more general than those in Bierens (1982a, 1984) and White and Domowitz (1984). White and Domowitz (1984) assume that the data generating process is α -mixing and they impose further conditions on the rate at which $\alpha(m)$ converges to zero as $m \rightarrow \infty$ in order to prove strong consistency and asymptotic normality of nonlinear least squares estimators. Bierens (1982a, 1984) assumes that the data generating process is stochastically stable with respect to a ϕ -mixing base. Here we shall assume that the data generating process is stochastically stable with respect to an α -mixing base without imposing further conditions on α .

This condition therefore encompasses the conditions of Bierens and White and Domowitz, as the ϕ -mixing condition implies the α -mixing condition and an α -mixing process is stochastically stable with respect to itself and hence to an α -mixing base. Moreover, we allow a fair amount of heterogeneity by assuming that the process is properly heterogeneous.

ASSUMPTION B1. The data generating process (z_t) with $z_t = (y_t, x_t')$ $\in \mathbb{R} \times \mathbb{R}^k$ is stochastically stable with respect to an α -mixing base and is properly heterogeneous. Moreover, $\sup_t E \|z_t\|^{4+\delta} < \infty$ for some $\delta > 0$.

(Cf. Definitions A1, A2 and A5).

In the sequel we shall denote the base involved by (u_t) where $u_t \in U$ with U a Euclidean space, and the mean process of (z_t) (cf. Definition A5) will be denoted by (z_t^*) . It should be noted that the error e_t of the ARMAX model (1.1) need not be a component of u_t , as it is possible that the e_t 's themselves are generated by a one-sided infinite sequence of u_t 's.

If we would make the strict stationarity assumption then Assumption B1 simplifies to:

ASSUMPTION B1*. There exist a strictly stationary α -mixing process (u_t) in a Euclidean space U and a Borel measurable mapping G from the space

of one-sided infinite sequences in U into R^{k+1} such that $z_t = (y_t, x_t)'$
 $= G(u_t, u_{t-1}, u_{t-2}, \dots)$ a.s. Moreover, $E\|z_t\|^{4+\delta} < \infty$ for some $\delta > 0$.

Thus:

THEOREM B1. Assumption B1* implies Assumption B1.

The proof of this theorem follows straightforwardly from Theorem A1, the fact that v -stability implies stochastic stability, and that by the strict stationarity assumption the proper heterogeneity condition automatically holds with mean process (z_t) . (Thus $z_t^* = z_t$ in this case).

Next consider the function $\bar{Q}(\theta)$ defined by (2.17). Let y_0^* be the first component of z_0^* and let

$$\bar{Q}(\theta) = E\{y_0^* - \phi(\theta) - \sum_{s=1}^{\infty} \eta_s(\theta)' z_{-s}^*\}^2$$

Then it follows from Theorem A3:

THEOREM B2: Under Assumption B1, $\text{plim}_{n \rightarrow \infty} \sup_{\theta \in \Theta} |\hat{Q}(\theta) - \bar{Q}(\theta)| = 0$.

Next, we assume

ASSUMPTION B2: There exists a unique $\theta_* \in \Theta$ such that $\bar{Q}(\theta_*) = \inf_{\theta \in \Theta} \bar{Q}(\theta)$.

Since Θ is compact and $\bar{Q}(\theta)$ is continuous there is always a θ_* in Θ which minimizes $\bar{Q}(\theta)$ over Θ . Thus the actual contents of this assumption is the uniqueness of θ_* . If the null hypothesis (1.2) is true then $\theta_* = \theta_0$, so that Assumption B2 then identifies the parameters of model (1.1). However, this assumption is also supposed to hold in the case that the null hypothesis (1.2) is false.

Applying Lemma 3.1.8. of Bierens(1981), it follows from Theorem B2 and Assumption B2:

THEOREM B3. Under Assumptions B1 and B2 the least squares estimator $\hat{\theta}$ defined by (2.18) satisfies $\text{plim}_{n \rightarrow \infty} \hat{\theta} = \theta_*$.

This proves (2.22).

Next we show the consistency and asymptotic normality of $\hat{\theta}$ under the assumption that model (1.1) is true in the sense that (1.2) holds for each t . Since (1.1) and (1.2) are equivalent to (2.8), we assume now:

ASSUMPTION B3. There exists a point θ_0 in an open subset Θ_0 of Θ , such that (2.8) holds for each t .

The condition that θ_0 lies in an open subset of Θ is hardly a restriction. The sets M , A , and B (cf. (2.4) and (2.6)) can be chosen to be the closures of open sets M_0 , A_0 and B_0 , respectively, whereas the set

$$\Gamma_{0,\delta} = \{(\gamma_1, \dots, \gamma_r)' \in R^r: |\lambda_s| < 1-\delta \text{ for } s=1,2,\dots,r\}$$

(cf. (2.3)) is for $\delta \in (0,1)$ the continuous image of an open set and therefore open itself, with closure Γ_δ . Thus $\Theta_0 = M_0 \times A_0 \times B_0 \times \Gamma_{0,\delta}$ (cf. (2.5)) is then an open set with closure Θ . In particular, if we let $M_0 = (-m,m)$, $A_0 = X_{\ell=1}^p(-m,m)$, $B_0 = X_{\ell=1}^{kq}(-m,m)$ and $\Theta_0 = M_0 \times A_0 \times B_0 \times \Gamma_{0,1/m}$, then Θ_0 is an open set with $\theta_0 \in \Theta_0$ for m sufficiently large and moreover its closure $\Theta = \bar{\Theta}_0$ is compact.

In order to establish the consistency of the least squares estimator $\hat{\theta}$ it suffices to show that θ_0 minimizes $\bar{Q}(\theta)$, as then by the uniqueness condition in Assumption B2, θ_* must be θ_0 . Showing this is easy and therefore left to the reader. Thus we have:

THEOREM B4. Under Assumption B1, B2 and B3, $\text{plim}_{n \rightarrow \infty} \hat{\theta} = \theta_0$.

Next, let

$$(B1) \quad g^*(\theta) = \phi(\theta) + \sum_{s=1}^{\infty} \eta_s(\theta)' z_{-s}^*,$$

where we recall that (z_t^*) is the mean process of (z_t) , and again let y_0^* be the first component of z_0^* . Denote

$$\Omega_1 = E\{(\partial/\partial\theta)' g^*(\theta_*)\} \{(\partial/\partial\theta) g^*(\theta_*)\},$$

and

$$\Omega_2 = E\{y_0^* - g^*(\theta_*)\}^2 \{(\partial/\partial\theta')g^*(\theta_*)\} \{(\partial/\partial\theta)g^*(\theta_*)\},$$

and assume:

ASSUMPTION B4. The matrix Ω_1 is non-singular.

This assumption is part of the set of maintained hypotheses which are assumed to hold regardless whether or not the null hypothesis is true.

THEOREM B5. Under Assumptions B1, B2, B3 and B4,

$$(B2) \quad \sqrt{n}(\hat{\theta} - \theta_0) \rightarrow N_m(0, \Omega_1^{-1} \Omega_2 \hat{\Omega}_1^{-1}) \text{ in distr.}$$

and under Assumptions B1, B2 and B4,

$$(B3) \quad \text{plim}_{n \rightarrow \infty} \hat{\Omega}_1^{-1} \hat{\Omega}_2 \hat{\Omega}_1^{-1} = \Omega_1^{-1} \Omega_2 \Omega_1^{-1}.$$

Note that Assumption B3 is not needed for (B3), due to the fact that in defining Ω_1 and Ω_2 we have used θ_* instead of θ_0 . However, if Assumption B3 does not hold then $\hat{\Omega}_1^{-1} \hat{\Omega}_2 \hat{\Omega}_1^{-1}$ is no longer the variance matrix of the limiting distribution of $\hat{\theta}$. Moreover, note that non-singularity of Ω_2 is not required: if Ω_2 is singular the limiting normal distribution (B2) is singular too.

The proof of Theorem B5 is not too hard and therefore left to the reader. The details can be found in the extended version of this appendix, which is available from the author on request.

APPENDIX C: THE TEST

In this appendix we prove Theorem 1 and we set forth the conditions for which the results in Section 2.3 hold. In the first instance it will be assumed that the data generating process (z_t) is strictly stationary. At the end we shall pay attention to the heterogenous case.

Proof of Theorem 1

Due to Theorem A1, the conditions (I) and (II) in Theorem 1 imply that the process $\{(v_t, z_t)\}$ is v -stable in L^1 with respect to the α -mixing base (u_t) . Moreover, since rounding off a real variable to ℓ -decimal digits defines a Borel measurable real function, the same applies to the process $\{(v_t, z_t^{(\ell)})\}$. The latter will enable us to apply Bierens (1986, Theorem 4).

Now let $F_{-\infty}^{t-1}$ be the Borel field generated by $z_{t-1}^{(1)}, z_{t-2}^{(2)}, z_{t-3}^{(3)}, \dots$ and let $F_{-\infty}^{t-1}(\ell)$ the Borel field generated by $z_{t-1}^{(\ell)}, z_{t-2}^{(\ell)}, z_{t-3}^{(\ell)}, \dots$. Then

$$(C1) \quad F_{-\infty}^{t-1}(\ell) \subset F_{-\infty}^{t-1}(\ell+1) \subset \dots \subset F_{-\infty}^{t-1}.$$

It follows therefore from Chung (1974, Theorem 9.4.8) that

$$(C2) \quad \lim_{\ell \rightarrow \infty} E(v_t | F_{-\infty}^{t-1}(\ell)) = E(v_t | F_{-\infty}^{t-1}) \quad \text{a.s.}$$

and consequently

$$(C3) \quad \lim_{\ell \rightarrow \infty} P\{E(v_t | z_{t-1}^{(\ell)}, z_{t-2}^{(\ell)}, \dots) = 0\} \\ = P\{E(v_t | z_{t-1}, z_{t-2}, \dots) = 0\}.$$

Because of stationarity, (C3) holds uniformly in t .

In Bierens(1986, Theorem 4) we have shown that, due to the rationality of $(z_t^{(\ell)})$, there exists a subset S_ℓ^* of $R^m \times (-1, 1)$ (possibly depending on ℓ) with Lebesgue measure zero, such that $(\xi, \tau) \in R^m \times (-1, 1) - S_\ell^*$ implies

$$(C4) \quad E(v_t | z_{t-1}^{(\ell)}, z_{t-2}^{(\ell)}, z_{t-3}^{(\ell)}, \dots) = E(v_t | \sum_{j=1}^{\infty} \tau^{j-1} \xi, z_{t-j}^{(\ell)}) \quad \text{a.s.}$$

Taking

$$(C5) \quad S^* = \bigcup_{\ell} S_{\ell}^*,$$

which is also a set with zero Lebesgue measure, and combining (C3) and (C4), we conclude that there exists a natural number ℓ_0 such that

$$(C6) \quad P\{E(v_t | z_{t-1}, z_{t-2}, \dots) = 0\} < 1$$

if and only if for each $\ell \geq \ell_0$ and each $(\xi, \tau) \in R^m \times (-1, 1) - S^*$,

$$(C7) \quad P\{E(v_t | \sum_{j=1}^{\infty} \tau^{j-1} \xi, z_{t-j}^{(\ell)}) = 0\} < 1.$$

Now let ψ be an arbitrary bounded Borel measurable one-to-one mapping. Then

$$(C8) \quad E(v_t | \sum_{j=1}^{\infty} \tau^{j-1} \xi, z_{t-j}^{(\ell)}) = E(v_t | \psi(\sum_{j=1}^{\infty} \tau^{j-1} \xi, z_{t-j}^{(\ell)})) \quad \text{a.s.,}$$

due to the fact that $\psi(\omega)$ with ω a random variable generates the same Borel field as ω . Since the conditioning variable on the right hand side of (C8) is bounded, the following lemma applies.

LEMMA C1. Let v be a random variable with $E|v| < \infty$ and let x be a bounded random variable. Then

$$(C9) \quad P(E(v|x) = 0) < 1$$

if and only if there is a countable subset C of R such that for each $\rho \in R - C$,

$$(C10) \quad E v \cdot \exp(\rho x) \neq 0.$$

Proof: From the proof of Theorem 2 of Bierens(1982b) it follows that there exists a $\delta > 0$ such that (C9) implies

$$(C11) \quad E v \cdot \exp(\rho x) \neq 0 \text{ for } \rho \in (-\delta, 0) \cup (0, \delta).$$

Now let for some $\rho_0 \in R$,

$$(C12) \quad E v \cdot \exp(\rho_0 x) = 0.$$

Since (C9) implies $P(E(v \cdot \exp(\rho_0 x) | x) \neq 0) < 1$, we may replace v in (C11) by $v \cdot \exp(\rho_0 x)$, which yields the conclusion that for some $\delta_0 > 0$, possibly depending on ρ_0 ,

$$(C13) \quad E v \cdot \exp(\rho x) \neq 0 \quad \text{for } \rho \in (\rho_0 - \delta_0, \rho_0) \cup (\rho_0, \rho_0 + \delta_0).$$

Now let C be the set of all these ρ_0 's. Then (C13) implies for each $\rho_0 \in C$

$$(C14) \quad \inf_{\rho \in C - \{\rho_0\}} |\rho - \rho_0| > 0.$$

and the same applies to any closed and bounded subset $\bar{C}$ of C containing at least two elements. However, if $\bar{C}$ contains infinitely many elements then it contains at least one limit point, which would violate (C14). Thus $\bar{C}$ is finite. Since C can be constructed as a countable union of such sets $\bar{C}$, C must be countable. Q.E.D.

Let us return to the proof of Theorem 1. Lemma C1 says that if

$$(C15) \quad P\{E(v_t | \psi(\sum_{j=1}^{\infty} \tau^{j-1} \xi, z_{t-j}^{(l)})) = 0\} < 1,$$

there exists a countable subset $C(\xi, \tau)$ of R , possibly depending on ξ and τ , such that for all $\rho \in R - C(\xi, \tau)$,

$$(C16) \quad E v_t \exp \{ \rho \psi(\sum_{j=1}^{\infty} \tau^{j-1} \xi, z_{t-j}^{(l)}) \} \neq 0.$$

Next we construct the subset S of $R \times R^m \times (-1, 1)$. Let for convenience

$$(C17) \quad I_l^*(\rho, \xi, \tau) = 1(E v_t \exp \{ \rho \psi(\sum_{j=1}^{\infty} \tau^{j-1} \xi, z_{t-j}^{(l)}) \} = 0),$$

where $I(\cdot)$ is the indicator function. Then (C6) implies that for each $(\xi, \tau) \in R^m \times (-1, 1) - S^*$, each $\rho \in R - C(\xi, \tau)$ and each $l \geq l_0$,

$$(C18) \quad I_{\ell}^*(\rho, \xi, \tau) = 0.$$

Now draw $\tilde{\rho}$ and the components of $\tilde{\xi}$ independently from the uniform $(-k, k)$ distribution and draw $\tilde{\tau}$ independently from the uniform $(-1, 1)$ distribution. Then for $(\xi, \tau) \notin S^*$ we have $E\{I(\tilde{\rho} \in C(\xi, \tau))\}$, hence

$$(C19) \quad E\{I_{\ell}^*(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})\} = E\{I_{\ell}^*(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})\}I(\tilde{\rho} \notin C(\tilde{\xi}, \tilde{\tau})) = 0,$$

where the last conclusion follows from (C18). Moreover, since $E\{I((\tilde{\xi}, \tilde{\tau}) \in S^*)\} = 0$ we have

$$(C20) \quad E\{I_{\ell}^*(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})\} = E\{I_{\ell}^*(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})\}I((\tilde{\xi}, \tilde{\tau}) \notin S^*).$$

Combining (C19) and (C20) yields

$$(C21) \quad E\{I_{\ell}^*(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})\} = 0.$$

Consequently, the set of $(\rho, \xi, \tau) \in X_{i=1}^{m+1}(-k, k) \times (-1, 1)$ for which $I_{\ell}^*(\rho, \xi, \tau) > 0$ must be a null set with respect of a Lebesgue measure. This null set, say $S_{k, \ell}$, depends on k and ℓ . Taking $S = \bigcup_k \bigcup_{\ell} S_{k, \ell}$ we now have a set with the required properties. This completes the proof of Theorem 1.

Q.E.D.

The test

Let V be a compact subset of $R^{k+1} \times (-1, 1)$ such that

$$(C22) \quad \inf_{(\xi, \tau) \in V} \{E w_{t, \ell}(\tau, \xi)^2 - (E w_{t, \ell}(\tau, \xi))^2\} > 0,$$

where $w_{t, \ell}(\tau, \xi)$ is defined by (2.54), and let

$$(C23) \quad W = [\rho_1, \rho_2] \times V, \text{ with } -\infty < \rho_1 < \rho_2 < \infty.$$

Moreover, let

$$(C24) \quad \inf_{(\rho, \xi, \tau) \in W} \bar{s}_{\ell}^2(\rho, \xi, \tau) > 0,$$

where $\bar{s}_\ell^2(\rho, \xi, \tau)$ is defined by (2.42). Then it follows from Theorems 1 and A3:

THEOREM C1. Let Assumptions B1*, B2 and B4 hold but not Assumption B3. Let $(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})$ be drawn randomly from the uniform distribution on W , where W defined by (C23) is such that (C22) and (C24) hold for sufficiently large ℓ . There exists a natural number ℓ_0 such that for each $\ell \geq \ell_0$, $\text{plim}_{n \rightarrow \infty} |\sqrt{n\hat{T}_\ell}(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})| = \infty$.

This theorem establishes the consistency of our test.

The conditions for asymptotic normality of our test statistic are:

THEOREM C2. Let W and $(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})$ be as in Theorem C1 and let Assumptions B1*, B2, B3 and B4 hold. Then $\sqrt{n\hat{T}_\ell}(\tilde{\rho}, \tilde{\xi}, \tilde{\tau}) \rightarrow N(0,1)$ in distribution.

For proving this theorem it suffices to prove

$$(C25) \quad \sqrt{n\hat{T}_\ell}(\rho, \xi, \tau) \rightarrow N(0,1) \text{ for each } (\rho, \xi, \tau) \in W,$$

as then for each $\kappa \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} E \exp\{i\kappa \sqrt{n\hat{T}_\ell}(\rho, \xi, \tau)\} = \exp(-\frac{1}{2}\kappa^2)$$

and hence by bounded convergence.

$$\begin{aligned} \lim_{n \rightarrow \infty} E \exp\{i\kappa \sqrt{n\hat{T}_\ell}(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})\} &= \lim_{n \rightarrow \infty} E\{E(\exp\{i\kappa \sqrt{n\hat{T}_\ell}(\tilde{\rho}, \tilde{\xi}, \tilde{\tau})\} | \tilde{\rho}, \tilde{\xi}, \tilde{\tau})\} \\ &= \exp(-\frac{1}{2}\kappa^2). \end{aligned}$$

We can now prove (C25) by showing first that (2.38) holds and then that the martingale central limit theorem in Bierens(1984) is applicable. The details of this proof and that of Theorem C1 are given in the extended version of this appendix, which is available from the author on request.

Heterogeneity

Until so far we have assumed strict stationarity. Most of our results, however, go through for the heterogenous case as well simply by replacing Assumption B1* with Assumption B1. The alternative hypothesis of our test then becomes:

$$H_1^*: P\{E(y_0^* - g^*(\theta_*) | z_{-1}^*, z_{-2}^*, \dots) = 0\} < 1,$$

where (z_t^*) is the mean process of $(z_t) = ((y_t, x_t'))$, y_0^* is the first component of z_0^* , $g^*(\theta)$ is defined by (B1) and θ_* is defined by Assumption B2. The test is then only consistent with respect to this alternative hypothesis and no longer with respect to the general alternative that the null is false. The reason for this is that if (1.2) fails to hold the true model may no longer be time invariant. See Bierens (1984, Sec.9) for a similar case.

REFERENCES

- Barro, R.J.: "Unanticipated Money Growth and Unemployment in the United States", The American Economic Review 67 (1977), 101-115.
- Barro, R.J.: "Unanticipated Money, Output, and the Price Level in the United States", Journal of Political Economy 86 (1978), 549-580.
- Barro, R.J.: "Unanticipated Money Growth and Unemployment in the United States: Reply", The American Economic Review 69 (1979), 1004-1009.
- Barro, R.J. and M. Rush: "Unanticipated Money and Economic Activity", in: Rational Expectations and Economic Policy, Ed. by S. Fischer, Chicago: The University of Chicago Press, 1980.
- Baumol, W.J.: Business Behaviour, Value and Growth, New York: Macmillan, 1959.
- Bierens, H.J.: Robust Methods and Asymptotic Theory in Nonlinear Econometrics, Lecture Notes in Economics and Mathematical Systems, vol. 192, Berlin-Heidelberg-New York: Springer Verlag, 1981.
- Bierens, H.J.: "A Uniform Weak Law of Large Numbers under φ -Mixing with Application to Nonlinear Least Squares Estimation", Statistica Neerlandica 36 (1982a), 81-86.
- Bierens, H.J.: "Consistent Model Specification Tests", Journal of Econometrics 20 (1982b), 105-134.
- Bierens, H.J.: "Uniform Consistency of Kernel Estimators of a Regression Function under Generalized Conditions", Journal of the American Statistical Association 77 (1983), 699-707.
- Bierens, H.J.: "Model Specification Testing of Time Series Regressions", Journal of Econometrics 26 (1984), 323-353.
- Bierens, H.J.: "ARMA Memory Index Modeling of Economic Time Series", 1986, forthcoming in Econometric Theory.
- Bierens, H.J. and J. Hartog: "On Specifying, Estimating and Testing Regression Models with Discrete Explanatory Variables, with an Empirical Application to the Earnings Function", Research Memorandum no. 8405, Department of Economics, University of Amsterdam, 1984.
- Brown, B.M.: "Martingale Central Limit Theorems", Annals of Mathematical Statistics 42 (1971), 59-66.
- Carlson, J.A.: "A Study of Price Forecasts", Annals of Economic and Social Measurement 6 (1977), 27-56.

- CBS: Maandschrift, 's-Gravenhage: Centraal Bureau voor de Statistiek, 1960-1983.
- Chung, K.L.: A Course in Probability Theory, New York-London: Academic Press, 1974.
- Doan, T., P.Litterman and C.Sims: "Forecasting and Conditional Projection Using Realistic Prior Distributions", Econometric Reviews 3 (1984), 1-100.
- Don, F.J.H.: "Constrained Nonlinear Least Squares", 1985, Forthcoming in Statistica Neerlandica.
- Feller, W.: "An Introduction to Probability Theory and its Applications, Vol. II", New York: John Wiley, 1966.
- Figlewski, S. and P.Wachtel: "The Formation of Inflationary Expectations", The Review of Economics and Statistics 63 (1981), 1-10.
- Friedman, M.: "The Role of Monetary Policy", The American Economic Review 58 (1968), 1-17.
- Granger, C.W.J.: "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", Econometrica 37 (1969), 424-438.
- Hannan, E.J., W.T.M.Dunsmuir and M.Deistler: "Estimation of Vector ARMAX Models", Journal of Multivariate Analysis 10 (1980), 275-295.
- Lucas, R.E.: "Econometric Testing of the Natural Rate Hypothesis", in: The Econometrics of Price Determination, Ed. by O.Eckstein, Washington D.C., 1972.
- Lucas, R.E.: "Some International Evidence on Output-Inflation Trade-Offs", The American Economic Review 63 (1973), 326-334.
- McLeish, D.L.: "A Maximal Inequality and Dependent Strong Laws", The Annals of Probability 3, (1975), 829-839.
- Mullineaux, D.J.: "On Testing for Rationality: Another Look at the Livingston Price Expectations Data", Journal of Political Economy 86 (1978), 329-336.
- Muth, J.F.: "Rational Expectations and the Theory of Price Movements", Econometrica 29 (1961), 315-335.
- Pesando, J.E.: "A Note on the Rationality of the Livingston Price Expectations", Journal of Political Economy 83 (1975), 849-858.
- Pesaran, M.H.: "A Critique of the Proposed Tests of the Natural Rate-Rational Expectations Hypothesis", The Economic Journal 92 (1982), 529-554.

- Pesaran, M.H. and A.S.Deaton: "Testing Non-Nested Nonlinear Regression Models", Econometrica 46 (1978), 677-694.
- Phelps, E.S.: "Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time", Economica 34 (1967), 254-281.
- Royden, H.L.: "Real Analysis", London, Macmillan, 1968.
- Sargent, T.J.: "Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment", Brookings Papers on Economic Activity 2 (1973), 429-472.
- Sargent, T.J.: "A Classical Macroeconometric Model for the United States", Journal of Political Economy 84 (1976), 207-237.
- Sargent T.J. and C.A.Sims: "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory", in: New Methods in Business Cycle Research: Proceedings from a Conference, Ed. by C.A.Sims, Minneapolis: Federal Reserve Bank of Minneapolis, 1977.
- Sargent, T.J. and N.Wallace: "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule", Journal of Political Economy 83 (1975), 241-254.
- Sims, C.A.: "Macroeconomics and Reality", Econometrica 48 (1980), 1-48.
- Sims, C.A.: "An autogressive Index Model for the U.S., 1948-1975", in: Large-Scale Macro-Econometric Models, Ed. by J.Kmenta and J.B.Ramsey, Amsterdam: North-Holland, 1981.
- White, H. and I.Domowitz: "Nonlinear Regression with Dependent Observations", Econometrica 52 (1984), 143-161.