Tilburg University

Sensitivity of Univariate AR(1) Time-series Forecasts Near the Unit Root

Banerjee, A.N.

Publication date:
1997

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Banerjee, A. N. (1997). *Sensitivity of Univariate AR(1) Time-series Forecasts Near the Unit Root*. (CentER Discussion Paper; Vol. 1997-88). Econometrics.

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.

Discussion paper

90

241

V sensitivity analysis
t time series
t forecasting
t unit root

VC20
VC22
VC53
VC13

Center
for
Economic Research

No. 9788

**SENSITIVITY OF UNIVARIATE AR(1) TIME-
SERIES FORECASTS NEAR THE UNIT ROOT**

By Anurag N. Banerjee

October 1997

ISSN 0924-7815

Sensitivity of univariate AR(1) time-series
forecasts near the unit root.¹

by

Anurag N Banerjee

CentER for Economic Research,

Tilburg University

The Netherlands

October 7, 1997

¹I am indebted to Jan Magnus for his comments, discussions and help in writing this paper. I am also grateful for the useful comments from Harald Uhlig and Parantap Basu. Constructive comments by the referee is also appreciated.

Abstract

We consider the linear time-series model $y_t = d_t + u_t$ ($t = 1, \dots, n$), where d_t is the deterministic trend, and u_t the stochastic term which follows an AR(1) process, $u_t = \theta u_{t-1} + \varepsilon_t$ with normal innovations ε_t . Various assumptions about the start-up will be made. Our main interest lies in the behaviour of the l - period ahead forecast $\hat{y}_{n+l}$ near $\theta = 1$. Unlike in other studies of AR(1) unit root processes, we do not to ask the question whether $\theta = 1$ or not. Instead we ask what happens to the forecast estimate near and at $\theta = 1$. For this purpose we define the first and second order sensitivity measure of the forecast $\hat{y}_{n+l}$ near $\theta = 1$. This measures the sensitivity of the forecast at the unit root.

We consider two deterministic trends : $d_t = \beta_1$ and $d_t = \beta_1 + \beta_2 t$. The forecast will be the Best Linear Unbiased forecast. We show that, when $d_t = \beta_1$, the number of observations has no effect on forecast sensitivity. When the deterministic trend is linear, the sensitivity is zero.

We also develop a large sample procedure to measure the forecast sensitivity when we are uncertain whether or not to include the linear trend.

Keywords : Sensitivity, Time-series, Unit Root, Forecast

JEL Codes : C20,C22,C53,C13

1 Introduction

In their seminal article Nelson and Plosser (1982) found that of the fourteen US economic time-series they considered, thirteen are non-stationary using Dickey and Fuller's (1979) formal unit root tests (see also Fuller (1976)). Similarly, Messe and Singleton (1982) applied Dickey-Fuller's tests and found that various exchange rate series are non-stationary. Following these two papers, there has been a surge of interest in the non-stationarity of economic time-series in the theoretical and empirical time series literature. Most of these papers concentrated on the estimation of and the inference on "unit roots" in the time-series.

In this paper, we are not interested whether the series is non-stationary or not. Instead we shall ask the question, whether or not the non-stationarity of the time-series matters in medium term forecasting? That is, how sensitive are the medium term forecasts against non-stationary behavior? This is a focus question in Uhlig (1994) in a Bayesian context.

We consider the univariate time series observations $\{y_t : t = 1, \dots, n\}$, which we write as a sum of deterministic trend d_t and a stochastic term u_t , that is $y_t = d_t + u_t$. The deterministic trend component d_t is known up to a finite dimensional parameter vector β . The stochastic component u_t follows an AR(1) process, with normal innovations, $u_t = \theta u_{t-1} + \varepsilon_t$, $t = 2, \dots, n$. The process is either assumed to be a part of an infinite horizon stationary process, in which case the variance of the process is not defined at $\theta = 1$ (Assumption A1), or a fixed start-up process at $t = 1$, with variance of the initial disturbance being constant (Assumption A2). The innovations $\{\varepsilon_t, t = 1, \dots, n\}$, are distributed $iid N(0, \sigma^2)$, where $\sigma^2 > 0$.

Let the l - period ahead forecast at $\theta = 1$, be defined as the limit of the l - period ahead forecast $\hat{y}_{n+l}(\theta, A)$, when $\theta \rightarrow 1$, where $A = A1, A2$ respectively. In section 3 we define the s^{th} order sensitivity measures of $\hat{y}_{n+l}(\theta; A)$ at $\theta = 1$ as

$$\lambda_l^{(s)}(A) = \lim_{\theta \rightarrow 1} \frac{d^s \hat{y}_{n+l}(\theta; A)}{d\theta^s}, s = 1, 2 \text{ and } A = A1, A2.$$

In sections 4 and 5, we concentrate on two leading deterministic specifications, a constant ($d_t = \beta_1$) and a linear trend ($d_t = \beta_1 + t\beta_2$), respectively.

In section 4 our results indicate that, the sensitivity of forecast is independent of the frequency of data collected, when the deterministic component is a constant. We show, in section 5, that the first order sensitivity measure is zero, when the deterministic trend is linear. This implies that there is no difference between the forecasts of a stationary model and a non-stationary model, with a linear trend up to an order of $(1 - \theta)^2$. In section 6 we derive the large sample distributions of our sensitivity measure. We also give rules of thumb to assess the sensitivity of the forecasts for different d_t 's and also in the case when we do not know whether we have a linear trend or a constant as d_t .

After some concluding remarks, we provide the proofs of theorems in Appendix 1. Appendix 2, contains two lemmas useful in proving the theorems and also of independent interest.

2 Preliminaries

We consider the vector of time series observations $y = (y_1, \dots, y_t, \dots, y_n)'$, and we write y_t as a sum of deterministic trend d_t and a stochastic term u_t ,

$$y_t = d_t + u_t \quad (2.1)$$

The trend d_t is known up to a finite dimensional parameter vector β . The stochastic component u_t follows an AR(1) process, with normal innovations,

$$\begin{aligned} u_1 &= \frac{1}{\delta} \varepsilon_1 \\ u_t &= \theta u_{t-1} + \varepsilon_t, \quad t = 2, \dots, n \end{aligned} \quad (2.2)$$

with $\varepsilon_1, \dots, \varepsilon_n \sim iid N(0, \sigma^2)$, where $\sigma^2 > 0$.¹ Regarding δ we shall make one of the following two assumptions.

Assumption A1: (stationary process): $\delta = \sqrt{1 - \theta^2}$, $|\theta| < 1$.

Under assumption A1, the variance-covariance matrix of $u = (u_1, \dots, u_n)'$, is $\sigma^2 \Omega(\theta; A1)$, where

$$\Omega(\theta; A1)_{ij} = \frac{\theta^{|i-j|}}{1 - \theta^2}, \quad |\theta| < 1. \quad (2.3)$$

Assumption A2: (fixed start-up): δ is a constant ($\neq 0$).

If $\delta = 1$, we have the covariance matrix of Berenblut and Webb (1973). In the time series forecasting literature, similar assumptions have been made in different contexts by several authors (see Dufour and King (1991), Magnus and Rothenberg (1988), Magnus and Pesaran (1991) and Fuller and Hasza (1980)).

Under Assumption A2, the variance-covariance matrix of $u = (u_1, \dots, u_n)$, is $\sigma^2 \Omega(\theta; A2)$, where

$$\Omega(\theta; A2)_{ij} = \theta^{|i-j|} \left(\sum_{k=0}^{\min(i,j)-2} \theta^{2k} + \frac{1}{\delta^2} \theta^{2(\min(i,j)-1)} \right). \quad (2.4)$$

¹We shall maintain this “structural” specification as opposed to a “reduced form” specification in order keep the same deterministic component for all θ .

It is useful to note that $\Omega(\theta; A1)^{-1}$ and $\Omega(\theta; A2)^{-1}$ can be written as, (see Dufour and King (1991))

$$\Omega(\theta; A1)^{-1} = (1 - \theta)^2 I_n + \theta D' D + \theta(1 - \theta)(e_1 e_1' + e_n e_n') \quad (2.5)$$

and

$$\Omega(\theta; A2)^{-1} = \Omega(\theta; A1)^{-1} + \left(\delta^2 - (1 - \theta^2)\right) e_1 e_1' \quad (2.6)$$

where I_n is the $n \times n$ identity matrix, e_1 and e_n are $n \times 1$ vectors $e_1 = (1, 0, \dots, 0, 0)'$, $e_n = (0, 0, \dots, 0, 1)'$, and D is defined as the $n \times n$ matrix of first differences,

$$D = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

3 Sensitivity of the forecast

When θ is known, and the deterministic term is specified and known, the Minimum Mean Square Error l - period ahead forecast of the y_t process given by (2.1) and (2.2) is

$$\hat{y}_{n+l}(\theta; A) = d_{n+l} + \theta^l (y_n - d_n), \quad (3.1)$$

under assumption $A = A1, A2$ (Harvey (1990, 1993)).

Since $d_t = d_t(\beta)$ and the parameter vector β is unknown, we use the Maximum Likelihood Estimate (MLE) of β , $\hat{\beta}(\theta; A)$, based on the past values of y_t under assumption $A = A1, A2$, in (3.1). As $\Omega(\theta; A1)$ is not defined at $\theta = 1$, we define the l - period ahead forecast at $\theta = 1$ as

$$\hat{y}_{n+l}(A1) = \lim_{\theta \rightarrow 1} \hat{y}_{n+l}(\theta; A1),$$

whenever the limit exists.

Subsequently we define the s^{th} order sensitivity of the forecast $\hat{y}_{n+l}(\theta; A)$ at $\theta = 1$ as

$$\lambda_l^{(s)}(A) = \lim_{\theta \rightarrow 1} \frac{d^s \hat{y}_{n+l}(\theta; A)}{d\theta^s}, \quad s = 1, 2 \text{ and } A = A1, A2,$$

whenever the derivatives and the limit exist. Justification for our proposed definition of sensitivity can be motivated by developing $\hat{y}_{n+l}(\theta; A)$ as a (limiting) Taylor series expansion in the neighborhood of $\theta = 1$,

$$\hat{y}_{n+l}(\theta; A) = \hat{y}_{n+l}(A) + (1 - \theta) \lambda_l^{(1)}(A) + \frac{1}{2} (1 - \theta)^2 \lambda_l^{(2)}(A) + \dots \quad (3.2)$$

We would consider $\hat{y}_{n+l}(\theta; A)$ and $\hat{y}_{n+l}(A)$ to be almost equal if

$$\lambda_l^{(s)}(A) \approx 0, \quad (s = 1, 2) \quad (3.3)$$

If the statistic $\lambda_l^{(s)}(A)$, is near to zero we need not discriminate between stationary and non stationary (unit root) process, as far as forecasting is concerned. If $\lambda_l^{(1)}(A)$ is zero or "near to zero" but $\lambda_l^{(2)}(A)$ is not, then the order of error is $(1 - \theta)^2$.²

²This will be a case when we take our deterministic component as a linear trend, which we shall discuss in more detail in section 5.

Given that the sensitivity measures exist, and since we are interested in knowing how close are $\hat{y}_{n+l}(\theta; A)$ and $\hat{y}_{n+l}(A)$, we need to study how close $\lambda_l^{(s)}(A)$ ($s = 1, 2$) is to zero. The $\lambda_l^{(s)}(A)$ ($s = 1, 2$) will generally be random variables³. So it will be useful to study the following probabilities as a measure of “closeness” to zero,

$$\pi(\theta) = \Pr_{\theta} \left(\left| \lambda_l^{(s)}(A) \right| \leq c_{\alpha} \right), \quad s = 1, 2, \quad (3.4)$$

where $\Pr_{\theta}$ is the probability measure associated with the random variable $u \sim N(0, \sigma^2 \Omega(\theta; A))$ ($A = A1, A2$), and c_{α} is obtained from the equation,

$$\lim_{\theta \rightarrow 1} \Pr_{\theta} \left(\left| \lambda_l^{(s)}(A) \right| \leq c_{\alpha} \right) = \alpha, \quad 0 < \alpha < 1. \quad (3.5)$$

$\pi(\theta)$ is essentially a robustness function of the $\lambda_l^{(s)}(A)$ statistic, against various distributive specifications of the error process. The probabilities give an indication of how close to zero the sensitivity measures are. The greater the probability mass of the sensitivity measures around zero there is closer is the distance between $\hat{y}_{n+l}(\theta; A)$ and $\hat{y}_{n+l}(A)$. In order to have a sharper bound for the sensitivity we will choose a lower value of α (in this paper we chose $\alpha = 0.05$). Higher the value of $\pi(\theta)$, lower is the probability of sensitivity of $\lambda_l^{(s)}(A)$. In this sense the sensitivity of the relevant statistic decrease when $\pi(\theta)$ increases. We also notice that the probabilities $\pi(\theta)$, are invariant under a monotonic transformation of $\lambda_l^{(s)}(A)$.

A similar study had been done by Banerjee and Magnus (1996), on the sensitivity of the Ordinary Least Squares (OLS) estimates. In their paper Banerjee and Magnus (1996) also shows the relation between the DW test statistic (and other related statistic like the Dicky-Fuller statistic and Kings alternative DW statistic) and sensitivity of OLS variance estimates of linear models.

In the next two sections, we shall study the sensitivity of forecasts of two leading cases of deterministic components, i) a constant $d_t = \beta_1$ and ii) a linear trend $d_t = \beta_1 + t\beta_2$. In this study the π curves described in (3.4) and (3.5) will be a useful device. Extensions can be made to higher

³See Section 5 where $\lambda_l^{(1)}(A) = 0$, $A = A1, A2$, when the deterministic component is linear ($d_t = \beta_1 + t\beta_2$).

order polynomial trend or trends satisfying more general conditions. The computation will be more involved but the method is the same.

The two deterministic components chosen, a constant and a linear trend, have also been studied extensively in the context of unit root testing and forecasting literature; see J.H Stock (1994) for an extensive review. Furthermore, as we will see our results will be easier to interpret and more useful to practitioners as they often use the two specifications in applied work.

4 Forecast sensitivity with a constant as the deterministic component

We consider the model with a constant but unknown trend $d_t = \beta_1$. From equation (3.1) the l -period ahead forecast is

$$\hat{y}_{n+l}(\theta, A) = \hat{\beta}_1(\theta; A) + \theta^l (y_n - \hat{\beta}_1(\theta; A)), \quad A = A1, A2, \quad (4.1)$$

where

$$\hat{\beta}_1(\theta; A1) = \frac{(1-\theta) \sum_{t=2}^{n-1} y_t + y_1 + y_n}{(1-\theta)(n-2) + 2}, \quad |\theta| < 1, \quad (4.2)$$

and

$$\hat{\beta}_1(\theta; A2) = \frac{(1-\theta)^2 \sum_{t=2}^{n-1} y_t + (1-\theta)(y_n - \theta y_1) + \delta^2 y_1}{(1-\theta)^2(n-1) + \delta^2} \quad (4.3)$$

are the Generalised Least Squares (GLS) estimates of β_1 , under assumption A1 and A2 respectively. It has been shown that $\hat{y}_{n+l}(\theta; A)$ is the BLUP (Best Linear Unbiased Predictor) when $0 < \theta < 1$; see Goldberger (1962).

Theorem 1 *We have*

- a) $\hat{y}_{n+l}(A1) = y_n$,
 - b) $\lambda_l^{(1)}(A1) = \frac{l}{2}(y_n - y_1)$,
 - c) $\lambda_l^{(2)}(A1) = l \sum_{t=2}^{n-1} (y_t - \hat{\beta}_1(A1)) + (l-1)\lambda_l^{(1)}(A1)$,
- where $\hat{\beta}_1(A1) = \frac{y_n + y_1}{2}$.

Theorem 2 *We have*

- a) $\hat{y}_{n+l}(A2) = y_n$,
 - b) $\lambda_l^{(1)}(A2) = l(y_n - y_1)$,
 - c) $\lambda_l^{(2)}(A2) = l\left(l - 1 + \frac{2}{\delta^2}\right)(y_n - y_1)$,
- where if $u \sim N(0, \sigma^2 \Omega(\theta; A2))$,
- d) $(y_n - y_1)/\sigma \sim N\left(0, \sum_{t=1}^{n-1} \theta^{t-1} + \frac{1}{\delta^2} (1 - \theta^{(n-1)})^2\right)$.

An intuition to discriminate between stationary and unit root alternative is that in case of a unit root the final observation wanders away from the initial observation (there is no mean reversing process). So we expect the difference $y_n - y_1$, to be “large” under unit root process and “small” when

we have less than unity root. A similar idea has been investigated using the R/S statistic $\frac{n^{-\frac{1}{2}}(\max_t y_t - \min_t y_t)}{n^{-\frac{1}{2}} \sum_{t=2}^n (y_t - y_{t-1})^2}$ by Mandelbrot and Van Ness (1968), Mandelbrot (1975) and Mandelbrot and Taqqu (1979). Although the statistic was originally proposed for measuring the fractionally integrated model, it proved useful in an autoregressive context also. (Stock (1994)). We can thus relate our sensitivity measures $\lambda_l^{(1)}(A1)$, $\lambda_l^{(1)}(A2)$, and $\lambda_l^{(2)}(A2)$ to same idea through Theorems 1 b) and 2 b) and c).

The sensitivity measures also addresses a major question in the empirical macro and financial economics literature, as to whether or not the frequency of data matters when forecasting at unit roots? The answer is it does not matter for 1st order sensitivity whether we assume A1 or A2 (Theorems 1 b) and 2 b)). For second order sensitivity, it matters when we have a limiting stationary process but not when we have a fixed start-up process. The intuition behind the difference, is that in the second order after discounting for the first order non-stationarity (Theorem 1 c)) we are left with a near independent situation, the sensitivity being the sum of errors.

In the unit root testing literature this is not an uncommon phenomenon (Campbell and Perron (1990)). It has been observed that, for a given span, additional observations obtained using data sampled more frequently, lead only to marginal improvement in discrimination (see Shiller and Perron (1985) and Perron (1990)).

Our results have been consistent with such observations, in particular if one considers the data generating process to be a fixed start-up model, but also to some extent for a limiting stationary model.

Let us now investigate the distributions generated by our measures in order to study the probabilities $\pi(\theta)$, which was motivated in section 3.

We begin by observing that the $\lambda_l^{(s)}(A)$, $s = 1, 2$ and $A = A1, A2$ are linear in y . Specially note that $\lambda_l^{(2)}(A1)$ can be written as $w_l' y$ where,

$$w_l = l \left(-\frac{1}{2}(n-l-1), 1, \dots, 1, -\frac{1}{2}(n+l-3) \right)'.$$

The sensitivity measures are normally distributed, as they are linear combinations of normally distributed variables. The distribution of $\lambda_l^{(s)}(A2)$, $s = 1, 2$ is obtained from Theorem 2 b) and d) and Theorem 2 c) and d) respectively.

Theorem 3 We have for, $u \sim N(0, \sigma^2 \Omega(\theta; A1))$,

- a) $\lambda_l^{(1)}(A1)/\sigma \sim N\left(0, \frac{l^2(1-\theta^{n-1})}{2(1-\theta^2)}\right) (|\theta| < 1)$,
- b) $\lambda_l^{(2)}(A1)/\sigma \sim N(0, w_1' \Omega(\theta; A1) w_1) (|\theta| < 1)$,
- c) $n^{-\frac{1}{2}} \lambda_l^{(1)}(A1)/\sigma \rightarrow N\left(0, \frac{1}{4} l^2 \left(1 - \frac{1}{n}\right)\right)$ as $\theta \rightarrow 1$,
- d) $n^{-\frac{3}{2}} \lambda_l^{(2)}(A1)/\sigma \xrightarrow{P} N\left(0, \frac{1}{12} l^2 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n} + \frac{3}{n^2} (l-1)^2\right)\right)$ as $\theta \rightarrow 1$.

Using Theorem 3 and Theorem 2 d), we can compute the probabilities $\pi(\theta)$, as described in section 3. The fact that the innovation variance σ^2 is not known does not matter in computation of the curves since the probabilities computed are scale independent, which is obvious from (3.4) and (3.5). So, for assessing the sensitivity of our model we need not know the exact value σ^2 , but if we ask whether the forecast obtained for a particular realisation of y is sensitive or not, we need to estimate σ^2 .

We estimate σ^2 with the usual estimator,

$$\widehat{\sigma^2} = \frac{1}{n-1} \sum_{t=2}^n (y_t - y_{t-1})^2, \quad (4.4)$$

under assumption $A = A1, A2$.

Theorem 4 We have, as $\theta \rightarrow 1$,

$$\frac{1}{n-1} \widehat{u}(\theta; A)' \Omega(\theta; A)^{-1} \widehat{u}(\theta; A) \xrightarrow{P} \widehat{\sigma^2}$$

where $\widehat{u}(\theta; A) = y - i\widehat{\beta}_1(\theta; A)$, $A = A1, A2$.

A rule of thumb to indicate sensitivity will be,

$$\frac{(\lambda_l^{(s)}(A))^2}{\widehat{\sigma^2}} \geq c_\alpha, \quad s = 1, 2. \quad (4.5)$$

The forecast is “sensitive” if (4.5) holds; otherwise not. Here c_α is obtained from the equation

$$\lim_{\theta \rightarrow 1} \Pr_{\theta} \left(\frac{(\lambda_l^{(s)}(A))^2}{\widehat{\sigma^2}} \leq c_\alpha \right) = \alpha, \quad 0 < \alpha < 1. \quad (4.6)$$

Since $\lambda_l^{(s)}(A)$, $s = 1, 2$, $A = A1, A2$, are linear in y ,

$$\frac{(\lambda_l^{(s)}(A))^2}{\widehat{\sigma^2}}, \quad s = 1, 2, \quad A = A1, A2,$$

is distributed as a ratio of quadratic forms of normally distributed random variables. The probability distribution functions of such ratios of quadratic forms can be easily computed by Imhof's method. (Imhof (1961), Koerts and Abrahamse (1969)).

It is to be noted here that $(\lambda_l^{(s)}(A1))^2 / \widehat{\sigma^2}$, $s = 1, 2$ is not defined at $\theta = 1$, but the probability limit of the random variables exists, as $\theta \rightarrow 1$.

Theorem 5 *We have, for $u \sim N(0, \sigma^2 \Omega(\theta; A1))$, as $\theta \rightarrow 1$,*

$$\begin{aligned} a) & \frac{4}{l^2(n-1)^2} \frac{(\lambda_l^{(1)}(A1))^2}{\sigma^2} \xrightarrow{p} B, \\ b) & \frac{4}{l^2(n-1)} \frac{(\lambda_l^{(2)}(A1))^2}{\sigma^2} \xrightarrow{p} \frac{(\sum_{t=1}^{n-1} (t - \frac{1}{2}(n-l+1)) v_t)^2}{\sum_{t=1}^{n-1} v_t^2}, \end{aligned}$$

where $B \sim \text{Beta}(\frac{1}{2}, \frac{n-2}{2})$ and $v_t \sim N(0, 1)$, $t = 1, \dots, n-1$.

We can then compare the π curves generated under assumption of known σ and unknown σ . Figure 1, shows the $\pi(\theta)$ curves of $\lambda_l^{(1)}(A1)/\sigma$ and $(\lambda_l^{(1)}(A1))^2/\widehat{\sigma^2}$, with $\alpha = 0.05$ and $n = 25, 100$. Figure 2, shows the $\pi(\theta)$ curves of $\lambda_l^{(2)}(A1)/\sigma$ and $(\lambda_l^{(2)}(A1))^2/\widehat{\sigma^2}$, with $\alpha = 0.05$, $n = 25, 100$ and $l = 5$. Figure 3, shows the $\pi(\theta)$ curves of $\lambda_l^{(1)}(A2)/\sigma$ and $(\lambda_l^{(1)}(A2))^2/\widehat{\sigma^2}$, $\delta = 1$, with $\alpha = 0.05$ and $n = 25, 100$. From Theorem 2 a) and b), $\lambda_l^{(1)}(A2)$ and $\lambda_l^{(2)}(A2)$ are the same random variable upto a multiplicative constant, so their π curves will be the same.

Figures 1, 2, and 3 show that the π curves are decreasing with θ and both the curves (known σ and unknown σ) in all the figures are close to each other. This implies that 1) the sensitivity of forecast is low for stationary distributions of u , since the probability mass of $\lambda_l^{(1)}(A)$ around zero increases as θ is away from the unit root. 2) In small samples estimating σ^2 understates the level of sensitivity, but the difference is small. As n increases the sensitivity also increases with θ , particularly the first order sensitivity, $\lambda_l^{(1)}(A)$, under assumptions $A = A1, A2$.

Theorem 6 *Under assumption A1 or A2, we have*

$$\begin{aligned} a) & \lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} n^{-\frac{1}{2}} (y_n - y_1) / \sigma \xrightarrow{p} N(0, 1), \\ b) & \lim_{n \rightarrow \infty} n^{-\frac{1}{2}} (y_n - y_1) / \sigma \xrightarrow{p} 0, \text{ if } |\theta| < 1. \end{aligned}$$

First order sensitivity degenerates to zero for stationary error processes that is the measure will show no sensitivity when the true distribution is stationary.

Fig.1. First order sensitivity : $d_i = \beta_i$, assumption A1.

Fig.2. Second order sensitivity : $d_t = \beta_t$, assumption A1.

Fig.3. First order sensitivity : $d_i = \beta_i$, assumption A2.

5 Forecast sensitivity with linear deterministic trend

We consider the model with a linear trend $d_t = \beta_1 + t\beta_2$, with unknown coefficients β_1 and β_2 . From equation (3.1) the l -period ahead forecast is

$$\begin{aligned}\hat{y}_{n+l}(\theta, A) &= \hat{\beta}_1(\theta; A) + (n+l)\hat{\beta}_2(\theta; A) \\ &+ \theta^l \left(y_n - \hat{\beta}_1(\theta; A) - n\hat{\beta}_2(\theta; A) \right),\end{aligned}\quad (5.1)$$

where $\hat{\beta}_1(\theta; A)$ and $\hat{\beta}_2(\theta; A)$ are the GLS estimates of β_1 and β_2 , under assumption $A = A1, A2$.

It is known that $\hat{y}_{n+l}(\theta; A)$ is the BLUP (Best Linear Unbiased Predictor) when $0 < \theta < 1$ (Goldberger (1962)).

Theorem 7 *We have, for $A = A1, A2$,*

- a) $\hat{y}_{n+l}(A) = y_n + l \frac{(y_n - y_1)}{n-1}$,
 - b) $\lambda_l^{(1)}(A) = 0$,
 - c) $\lambda_l^{(2)}(A) = \frac{2l}{n-1} \sum_{t=2}^{n-1} (y_t - \hat{\beta}_1 - t\hat{\beta}_2)(t-1)$,
- where $\hat{\beta}_1 = \frac{1}{n-1}(ny_1 - y_n)$ and $\hat{\beta}_2 = \frac{1}{n-1}(y_n - y_1)$.

Theorem 7 b) makes an important observation about the linear trend model: *If there is a deterministic linear time trend in the model, there is no difference (in terms of first order sensitivity) between the forecast of a unit root process and a stationary process irrespective of whether they are generated by the stationary assumption (A1) or the fixed-start-up assumption (A2).*

It has been long observed in Bayesian and classical unit root testing literature that it is difficult to discriminate between a trend stationary and a unit root process. Campbell and Perron (1990) stated four rules (Rules 7-10) about the problem of discrimination. Rules 8 and 10 say that in finite samples any test for unit root hypothesis against trend stationary alternatives must have power no greater than its size, and vice versa. This means that there cannot be any discrimination between the two regimes in finite samples. Theorem 7 b), says that in case of linear time trend, the discrimination in the forecast estimates, if any, can only happen through a second order

measure of sensitivity of the same. Note, that, if we plot the π curves for $\lambda_l^{(1)}(A)$, it will be a straight horizontal line. This brings us to Campbell and Perron's Rules 7 and 8, which say that in finite samples any unit root process can be approximated arbitrarily well by a trend-stationary process, and vice versa. In case of linear trend, Theorem 7 tells us that the approximation of the forecast estimate is of the order $(1 - \theta)^2$ (see equation 3.2).

Does that mean we should give up the idea to discriminate between stationary and non-stationary processes, in case of linear trend models as suggested by Christiano and Eichenbaum (1990)? Our answer is that we should look at the second order sensitivity statistic $\lambda_l^{(2)}(A)$ as well. From part c) of Theorem 7 we can write $\lambda_l^{(2)}(A)$ as $w_2'y$, where

$$w_2 = \frac{2l}{n-1} \left(-\frac{1}{6}n(n-2), 1, \dots, t-1, \dots, n-1, -\frac{1}{6}(n-2)(2n-3) \right)'$$

Then we have the following result.

Theorem 8 *If $u \sim N(0, \sigma^2 \Omega(\theta; A))$, $A = A1, A2$, we have*

- a) $\lambda_l^{(2)}(A)/\sigma \sim N(0, w_2'\Omega(\theta; A)w_2) \mid \theta \mid < 1$.
- b) $n^{-\frac{3}{2}}\lambda_l^{(2)}(A)/\sigma \xrightarrow{P} N\left(0, \frac{4}{45}l^2\frac{n-2}{n-1}\left(1 - \frac{1}{2n}\right)\left(1 - \frac{3}{2n}\right)\right)$ as $\theta \rightarrow 1$.

Using Theorem 8 we can compute the probabilities $\pi(\theta)$, as described in Sections 3 and 4. We also need to estimate σ^2 in order to construct a similar rule of thumb as in (4.5) and (4.6). We estimate σ^2 with the usual estimator

$$\widehat{\sigma^2} = \frac{1}{n-2} \left(\sum_{t=2}^n (y_t - y_{t-1})^2 - \frac{(y_n - y_1)^2}{n-1} \right), \quad (5.2)$$

under assumption $A = A1, A2$.

Theorem 9 *We have, as $\theta \rightarrow 1$,*

$$\frac{1}{n-2} \widehat{u}(\theta; A)' \Omega(\theta; A)^{-1} \widehat{u}(\theta; A) \xrightarrow{P} \widehat{\sigma^2}$$

where $\widehat{u}(\theta; A) = y - i\widehat{\beta}_1(\theta; A) - \tau\widehat{\beta}_2(\theta; A)$, $A = A1, A2$ and $\tau = (1, \dots, n)'$.

The rule of thumb is,

$$\frac{\left(\lambda_l^{(2)}(A)\right)^2}{\widehat{\sigma^2}} \geq c_\alpha, \quad 0 < \alpha < 1. \quad (5.3)$$

The forecast is sensitive if (5.3) holds; otherwise not. Here c_α is obtained from the equation,

$$\lim_{\theta \rightarrow 1} \Pr_{\theta} \left(\frac{\left(\lambda_l^{(2)}(A) \right)^2}{\sigma^2} \leq c_\alpha \right) = \alpha, \quad 0 < \alpha < 1. \quad (5.4)$$

As $\left(\lambda_l^{(2)}(A1) \right)^2 / \widehat{\sigma^2}$, is not defined at $\theta = 1$, we compute the probability limit as $\theta \rightarrow 1$.

Theorem 10 *We have, if $u \sim N(0, \sigma^2 \Omega(\theta; A1))$, as $\theta \rightarrow 1$,*

$$\frac{9}{l^2} \frac{\left(\lambda_l^{(2)}(A1) \right)^2}{\sigma^2} \xrightarrow{p} \frac{\left(\sum_{t=1}^{n-1} (n(n-2) + 3t(t-1)) v_t \right)^2}{(n-1) \sum_{t=1}^{n-1} v_t^2 - \left(\sum_{t=1}^{n-1} v_t \right)^2}$$

where $v_t \sim N(0, 1)$, $t = 1, \dots, n-1$.

We can now compare the π curves generated under assumption of known σ and unknown σ . Figure 4 shows the $\pi(\theta)$ curves of $\lambda_l^{(2)}(A1)/\sigma$ and $\left(\lambda_l^{(2)}(A1) \right)^2 / \widehat{\sigma^2}$, with $\alpha = 0.05$ and $n = 25, 100$. Figure 5 shows the $\pi(\theta)$ curves of $\lambda_l^{(2)}(A2)/\sigma$ and $\left(\lambda_l^{(2)}(A2) \right)^2 / \widehat{\sigma^2}$, $\delta = 1$, with $\alpha = 0.05$ and $n = 25, 100$.

Figures 4 and 5 indicate that the π curves are very flat and that both curves (known σ and unknown σ) in the all the figures are close to each other. We also notice that the second order sensitivity is of stationary distributions are almost same as unit root distributions, even for 100 observations. Thus we may conclude that, the sensitivity measure of the forecast of a linear trend model is robust to non-stationarity of the AR(1) error process.

Fig.4. Second order sensitivity: $d_t = \beta_t + t\beta_2$, assumption A1.

Fig.5. Second order sensitivity: $d_i = \beta_i + t\beta_2$, assumption A2.

6 Large sample sensitivity with an unknown deterministic component

The discussion so far has assumed that the deterministic component can either be a constant or a linear trend, but is correctly specified. In empirical studies, however, we will not know what is the true deterministic component.

The analogy to the usual regression problem is that of an omitted variable bias. In the regression problem the bias vanishes asymptotically as $n \rightarrow \infty$, since asymptotically the regressors and the error process are uncorrelated, the order of convergence being $n^{-\frac{1}{2}}$. When the error process is a random walk, the omitted variable bias also vanishes asymptotically but the rates of convergence (in n) depends on the order of the trend polynomial. As the constant and the linear trend are special cases of polynomial trend regressors the convergence rates of the test statistics are n^{-1} and $n^{-\frac{3}{2}}$ respectively.

The above observations relate to Campbell and Perron's (1991) Rule 4, which says that if we omit the time trend term (that is, we assume incorrectly that $\beta_2 = 0$) and just include the constant as the deterministic component, then we cannot discriminate between a stationary case and an unit root case since the power of unit root tests goes to zero. On the other hand, Rule 2 and 5, of Campbell and Perron (1991) tell us that, if we assume that $\beta_2 \neq 0$ when in fact $\beta_2 = 0$, then also the power of discrimination between the regimes goes to zero. The reason being, when we have polynomial trends as regressors, the asymptotic distributions of the test statistics have different orders of (n) of consistency. See Perron (1988) for an extensive analysis.

In our problem, we are not interested in testing the hypothesis of unit roots, but we are interested to know how sensitive is the forecast $\hat{y}_{n+i}(\theta; A)$ near $\theta = 1$. Can we assess the sensitivity of the forecast if we do not know the deterministic component? Or do we face the same problem as the unit root testing literature. We will answer the question by comparing the asymptotic distributions of $\lambda_l^{(2)}(A) / \sqrt{\sigma^2}$, as $n \rightarrow \infty$. We state the following theorem.

Theorem 11 *If $d_t = \beta_1 + t\beta_2$ and $u \sim N(0, \sigma^2\Omega(\theta; A))$, then*

$$a) \lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} n^{-\frac{3}{2}} \lambda_l^{(2)}(A) / \sqrt{\sigma^2} \stackrel{P}{\rightarrow} N\left(0, \frac{4}{45} l^2\right), \quad A = A1, A2.$$

If $d_t = \beta_1$ and $u \sim N(0, \sigma^2\Omega(\theta; A1))$, then

- b) $\lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} n^{-\frac{1}{2}} \lambda_l^{(1)}(A1) / \sqrt{\widehat{\sigma^2}} \stackrel{P}{=} N\left(0, \frac{l^2}{4}\right),$
c) $\lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} n^{-\frac{3}{2}} \lambda_l^{(2)}(A1) / \sqrt{\widehat{\sigma^2}} \stackrel{P}{=} N\left(0, \frac{1}{12} l^2\right)$
If $d_t = \beta_1$ and $u \sim N(0, \sigma^2 \Omega(\theta; A2))$, then
d) $\lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} n^{-\frac{1}{2}} \lambda_l^{(1)}(A2) / \sqrt{\widehat{\sigma^2}} \stackrel{P}{=} N(0, l^2),$
e) $\lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} n^{-\frac{1}{2}} \lambda_l^{(2)}(A2) / \sqrt{\widehat{\sigma^2}} \stackrel{P}{=} N\left(0, l^2 \left(l - 1 + \frac{3}{l^2}\right)^2\right).$

The theorem provides us with mixed answers. In case of Assumption A1, a) and c) of Theorem 11 show that the convergence of the second order sensitivity measure, both in the case of linear trend and constant, is of the order $n^{-\frac{3}{2}}$. Also notice that the asymptotic variance $\lambda_l^{(2)}(A1) / \sqrt{\widehat{\sigma^2}}$, when $d_t = \beta_1 + t\beta_2$, is $\frac{4l^2}{45} = \frac{l^2}{11.25}$, which is very close to $\frac{l^2}{12}$, the asymptotic variance of $\lambda_l^{(2)}(A1) / \sqrt{\widehat{\sigma^2}}$, when $d_t = \beta_1$. Let us consider the asymptotic equivalent of the rule of thumb described in (4.5) and (4.6), and (5.3) and (5.4), for $\left(\lambda_l^{(2)}(A1)\right)^2 / \widehat{\sigma^2}$. The c_α obtained from (4.6) will be very close to the one obtained from (5.4), so that we can use the same decision rule for the second order sensitivity, $\lambda_l^{(2)}(A1) / \sqrt{\widehat{\sigma^2}}$. Hence irrespective whether $\beta_2 = 0$ or not, we assume a linear trend, estimate β_2 , and use our sensitivity measure as in the case of a linear trend (rule of thumb (5.3) and (5.4)). In case we want to measure the first order sensitivity, we know by Theorem 7 that $\lambda_l^{(1)}(A1) = 0$, so we assume $\beta_2 = 0$ and use our rule of thumb (4.5) and (4.6).

Under assumption A2, the convergence rate of the second order sensitivity measure $\lambda_l^{(2)}(A2) / \sqrt{\widehat{\sigma^2}}$ is $n^{-\frac{3}{2}}$ and $n^{-\frac{1}{2}}$, when $d_t = \beta_1 + t\beta_2$ and $d_t = \beta_1$ respectively. Therefore, we have the same problem as the unit root testing literature and cannot use the same procedure as above. Assumption A1 is therefore necessary for the sequential decision making described above, when we are unsure of including, a constant or a linear trend, as a deterministic component.

7 Conclusion

In this paper we have introduced sensitivity measures $\lambda_l^{(1)}(A)$ and $\lambda_l^{(2)}(A)$ for our analysis of medium term forecast sensitivity near unit root. We have investigated two leading cases in time series analysis namely the model with a linear deterministic trend, and the model with a constant. It will not be difficult to extend the analysis to higher order polynomial trends or trend break models, in a similar way.

Another extension of interest would be to specify innovations ε_t to be an ARMA process instead of white noise. Some simulation experiments have been done in this direction by Campbell and Perron (1991) and Cochrane (1991). They studied the forecast changes when the innovation is distributed as an MA(1) process but is assumed to be *iid*. They conclude that it might be beneficial for forecasting efforts to have such misspecifications. Our framework in this paper presents us with an opportunity, to extend the analysis to study such hypotheses, by using the relevant π curves.

In section 6, we measure the second order sensitivity of the forecast assuming a linear deterministic component, when we are unsure whether or not to include the time trend component. We then assume a constant deterministic component to measure the first order sensitivity. We can therefore ask, whether a similar analysis can be done starting with higher order polynomials, by introducing higher order sensitivity measures

$$\lambda_l^{(s)}(A) = \lim_{\theta \rightarrow 1} \frac{d^s \hat{y}_{n+l}(\theta; A)}{d\theta^s}, s > 2 \text{ and } A = A1, A2.$$

Further work is also necessary to propose a method of correction in case the forecast turns out to be sensitive.

8 Appendix 1: Proofs of theorems

Proof of Theorem 1:

- a) As $\theta \rightarrow 1$, from equation (4.1) and (4.2) we get $\hat{y}_{n+l}(A1) = y_n$.
 b) Differentiating (4.1) and taking limits as $\theta \rightarrow 1$, we get: $\lambda_l^{(1)}(A1) = l(y_n - \lim_{\theta \rightarrow 1} \hat{\beta}_1(\theta; A1))$. Also $\lim_{\theta \rightarrow 1} \hat{\beta}_1(\theta; A1) = (y_n + y_1)/2$.
 c) Differentiating (4.1) twice and taking limits as $\theta \rightarrow 1$, we get

$$\lambda_l^{(1)}(A1) = l(l-1) \left(\frac{y_n - y_1}{2} \right) - 2l \lim_{\theta \rightarrow 1} \frac{d\hat{\beta}_1(\theta; A1)}{d\theta}$$

and we have

$$\lim_{\theta \rightarrow 1} \frac{d\hat{\beta}_1(\theta; A1)}{d\theta} = - \sum_{t=1}^{n-1} (y_t - \hat{\beta}_1(A1)).$$

Proof of Theorem 2:

- a) We have from (4.1) and (4.2) at $\theta = 1$, $\hat{y}_{n+l}(A2) = y_n$.
 b) From equation (4.1), differentiating (4.2) at $\theta = 1$, we have $\lambda_l^{(1)}(A2) = l(y_n - \hat{\beta}_1(1; A2))$ and $\hat{\beta}_1(1; A2) = y_1$.
 c) Differentiating (4.1) at $\theta = 1$,
 $\lambda_l^{(2)}(A2) = l(l-1)(y_n - y_1) - 2l \left. \frac{d\hat{\beta}_1(\theta; A2)}{d\theta} \right|_{\theta=1}$ and $\left. \frac{d\hat{\beta}_1(\theta; A2)}{d\theta} \right|_{\theta=1} = \delta^2(y_n - y_1)$.
 d) We have $Var(u_1 - u_n) = \sigma^2 \left(\sum_{t=1}^n \theta^{t-1} + \delta^2 (1 - \theta^{(n-1)})^2 \right)$.

Proof of Theorem 3:

- a) i) and ii) follow directly from the fact $u \sim N(0, \sigma^2 \Omega(\theta; A1))$.
 b) i) follows directly from a) i) and ii) follows from Lemma 1. Note that $w'_1 i = 0$ and u is a normally distributed. Using standard algebraic identities gives the result.

Proof of Theorem 4:

Under assumption A1, the result follows from (2.5) and the fact that $D\hat{z} = 0$.
 Under assumption A2, the result follows from (2.6) and $\hat{\beta}_1(1; A2) = y_1$.

Proof of Theorem 5:

- a) In Lemma 2, let $B = D'D$ and $w = \frac{1}{2}(1, \dots, -1)'$, and we get $\bar{w} = \frac{1}{2}(0, \dots, -1)'$ and $P'D'DP = I_{n-1}$, and it follows that the limit is $\frac{1}{4}(n-1)^2 \frac{(\sum_{t=1}^{n-1} v_t)^2 / (n-1)}{\sum_{t=1}^{n-1} v_t^2}$. Using the fact that $v_1, \dots, v_{n-1} \sim N(0, 1)$, we get the result.
 b) In Lemma 2, let $B = D'D$ and $w = w_1$, and observe that $i'w_1 = 0$.

Proof of Theorem 6:

a) Follows from Theorem 2 d) and Theorem 3 c).

b) Follows from Theorem 2 d) and Theorem 3 a).

Proof of Theorem 7:

We can write $\hat{\beta}_1(\theta; A)$ and $\hat{\beta}_2(\theta; A)$ as,

$$\hat{\beta}_1(\theta; A) = \frac{(\tau'\Omega^{-1}\tau)(i'\Omega^{-1}y) - (i'\Omega^{-1}\tau)(\tau'\Omega^{-1}y)}{(\tau'\Omega^{-1}\tau)(i'\Omega^{-1}i) - (i'\Omega^{-1}\tau)^2} \quad (8.1)$$

$$\hat{\beta}_2(\theta; A) = \frac{(i'\Omega^{-1}i)(\tau'\Omega^{-1}y) - (i'\Omega^{-1}\tau)(\tau'\Omega^{-1}y)}{(\tau'\Omega^{-1}\tau)(i'\Omega^{-1}i) - (i'\Omega^{-1}\tau)^2}, \quad (8.2)$$

where i is a $n \times 1$ vector of ones, and $\tau = (1, \dots, n)'$, and $\Omega^{-1} = \Omega(\theta; A)^{-1}$, $A = A1, A2$.

Observe that if $a = (a_1, \dots, a_t, \dots, a_n)'$ and $b = (b_1, \dots, b_t, \dots, b_n)'$, are any two vectors from (2.5), we have

$$\begin{aligned} a'\Omega(\theta; A1)^{-1}b &= (1-\theta)^2 \sum_{t=2}^{n-1} a_t b_t + \theta \sum_{t=2}^n (a_t - a_{t-1})(b_t - b_{t-1}) \\ &\quad + (1-\theta)(a_1 b_1 + a_n b_n) \end{aligned} \quad (8.3)$$

and from (2.6),

$$a'\Omega(\theta; A2)^{-1}b = a'\Omega(\theta; A1)^{-1}b + (\delta^2 - (1-\theta)^2) a_1 b_1. \quad (8.4)$$

Since $Di = 0$ the scalar $(1-\theta)$ factors out from both numerator and denominator of $\hat{\beta}_1(\theta; A1)$ and $\hat{\beta}_2(\theta; A1)$ ⁴. Using (8.3) and (8.4) we have the following:

a) $\lim_{\theta \rightarrow 1} \hat{\beta}_1(\theta; A) = \hat{\beta}_1$, $\lim_{\theta \rightarrow 1} \hat{\beta}_2(\theta; A) = \hat{\beta}_2$, and $\hat{y}_{n+l}(A) = y_n + l\hat{\beta}_2$.

b) Differentiating (5.1) and taking limits as $\theta \rightarrow 1$, we get,

$\lambda_i^{(1)}(A) = l(d\hat{\beta}_2(A) + y_n - \hat{\beta}_1 - n\hat{\beta}_2)$, $A = A1, A2$ We also have from

(8.1), (8.2), $d\hat{\beta}_2(A) = \lim_{\theta \rightarrow 1} \frac{d\hat{\beta}_2(\theta; A)}{d\theta} = 0$, and $y_n = \hat{\beta}_1 + n\hat{\beta}_2$.

c) Differentiating (5.1) twice and taking limits as $\theta \rightarrow 1$, we get

$$\lambda_i^{(2)}(A) = l(d^2\hat{\beta}_2(A) - 2d\hat{\beta}_1(A)), A = A1, A2,$$

⁴The cross product term $\sum_{t=2}^n (a_t - a_{t-1})(b_t - b_{t-1})$ vanishes if either a_t or $b_t = 1$, for all t .

where

$$\begin{aligned}
d\hat{\beta}_1(A1) &= \lim_{\theta \rightarrow 1} \frac{d\hat{\beta}_1(\theta; A1)}{d\theta} = -\frac{1}{2} \sum_{t=1}^{n-1} (y_t - \hat{\beta}_1 - t\hat{\beta}_2), \\
d^2\hat{\beta}_2(A1) &= \lim_{\theta \rightarrow 1} \frac{d^2\hat{\beta}_2(\theta; A1)}{d\theta^2} = \frac{2}{n-1} \sum_{t=1}^{n-1} y_t \left(t - \frac{n+1}{2} \right) - t^2 \hat{\beta}_2 \\
&\quad + \frac{n+1}{2} \hat{\beta}_1 - t \frac{(n+3)y_n - (1+3n)y_1}{2(n-1)}, \\
d\hat{\beta}_1(A2) &= \lim_{\theta \rightarrow 1} \frac{d\hat{\beta}_1(\theta; A2)}{d\theta} = 0, \\
d^2\hat{\beta}_2(A2) &= \lim_{\theta \rightarrow 1} \frac{d^2\hat{\beta}_2(\theta; A2)}{d\theta^2} = \frac{2}{n-1} \sum_{t=2}^{n-1} (y_t - \hat{\beta}_1 - t\hat{\beta}_2) (t-1).
\end{aligned}$$

Rewriting we get the result.

Proof of Theorem 8:

- a) Follows directly from the fact $u \sim N(0, \sigma^2 \Omega(\theta; A))$, $A = A1, A2$.
- b) follows from Lemma 1, noticing that $w_2' i = 0$ and u is a normally distributed. Using standard algebraic identities gives the result.

Proof of Theorem 9:

Under assumption A1, the result follows from (2.5) the fact that $Di = 0$.

Under assumption A2, the result follows from (2.6) and $y_1 = \hat{\beta}_1 + \hat{\beta}_2$.

Proof of Theorem 10:

- a) In Lemma 2, let $B = D'D - \frac{1}{n-1} D'D\tau\tau'D'D$, where $\tau = (1, \dots, n)'$, is a $n \times 1$ vector and $w = w_2$. We observe that $i'w_2 = 0$ and $P'\bar{B}P = I_{n-1} - \frac{1}{n-1} \bar{i}'\bar{i}$, where $\bar{i}$ is a $(n-1) \times 1$ vector of ones, and the result follows.

Proof of Theorem 11:

- a) If $d_t = \beta_1 + t\beta_2$ then from (5.2), $\lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} \widehat{\sigma^2} \stackrel{P}{=} \sigma^2$. Hence the result follows from Theorem 8 part c).

If $d_t = \beta_1$, then from (4.4), $\lim_{n \rightarrow \infty} \lim_{\theta \rightarrow 1} \widehat{\sigma^2} \stackrel{P}{=} \sigma^2$. Thus, b) follows from Theorem 1 part b) i), c) follows from Theorem 1 part b) ii), d) follows from Theorem 2 part b) and d), and e) follows from Theorem 2 part c) and d).

9 Appendix 2: Two useful lemmas

Lemma 1 Let $w = (w_1, \dots, w_n)'$. Then,

$$\lim_{\theta \rightarrow 1} w' \Omega(\theta; A1) w = \begin{cases} -\sum_{k=1}^{n-1} k \sum_{t=1}^{n-k} w_t w_{t+k} & ; \text{ if } w'i = 0 \\ \infty & ; \text{ otherwise} \end{cases}$$

where i is an $n \times 1$ vector of ones.

Proof of Lemma 1:

We can write

$$\Omega(\theta; A1) = \frac{1}{1 - \theta^2} \sum_{k=0}^{n-1} \theta^k T^{(k)}$$

where $T^{(k)}$ are Toeplitz matrices of order k (see Banerjee and Magnus (1996)). Since $\sum_{k=0}^{n-1} T^{(k)} = ii'$, $w' \Omega(\theta; A1) w$ at $\theta = 1$ takes a $\frac{0}{0}$ form if and only if $w'i = 0$. Hence we get the limit as

$$\lim_{\theta \rightarrow 1} w' \Omega(\theta; A1) w = -\frac{1}{2} \sum_{k=1}^{n-1} k w' T^{(k)} w,$$

otherwise it is infinity.

Lemma 2 Let $u \sim N(0, \sigma^2 \Omega(\theta; A1))$, $w = (w_1, \dots, w_n)'$, and $B \neq 0$, a symmetric positive semi-definite $n \times n$ matrix. Then, as $\theta \rightarrow 1$,

$$\frac{(w'u)^2}{u'Bu} = \begin{cases} \frac{(\bar{w}'Pv)^2}{v'P'BPv} & , \text{ if } w'i = 0, Bi = 0, \\ \infty & , \text{ if } w'i \neq 0, Bi = 0, \\ \frac{(w'i)^2}{i'Bi} & , \text{ if } Bi \neq 0, \end{cases}$$

where $\bar{w} = (w_2, \dots, w_n)'$, and $\bar{B}$ is the matrix obtained from B deleting the 1st row and the 1st column, $v \sim N(0, I_{n-1})$, P is a lower triangle $(n-1) \times (n-1)$ matrix such that,

$$P_{ij} = \begin{cases} 1 & \text{if } i \geq j \\ 0 & \text{if } i < j \end{cases}.$$

Proof of Lemma 2:

The proof directly follows from Theorem A1 of Banerjee and Magnus (1996).

10 References

- Banerjee, A.N. and J.R.Magnus, 1996, The sensitivity of OLS when the variance matrix is partially unknown, (submitted to *Journal of Econometrics*).
- Berenblut, I.I. and G.I. Webb, 1973, A new test for autocorrelated errors in the linear regression model, *Journal of the Royal Statistical Society*, B35, 33-50.
- Campbell, J.Y. and P.Perron, 1991, Pitfalls and opportunities: What macroeconomists should know about unit roots, in O.J.Blanchard and S.Fisher (eds.), *NBER Macroeconomics Annual*, (Cambridge: MIT Press), 141-201.
- Christiano, L.J. and M.Eichenbaum, 1990, Unit roots in GNP : Do we know, and do we care ?, *Carnegie-Rochester Conference Series on Public Policy*, 32, 7-62.
- Cochrane, J.H., 1991, Comment. in O.J.Blanchard and S.Fisher (eds.), *NBER Macroeconomics Annual*, (Cambridge: MIT Press), 201-210.
- Dickey, D.A. and W.A.Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-431.
- Dufour, J.-M. and M.L. King, 1991, Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors, *Journal of Econometrics*, 47, 115-143.
- Fuller, W.A., 1976, *Introduction to Statistical Time Series* (John Wiley, Chichester/ New York).
- Fuller, W.A. and D.P.Hasza, 1980, Predictors for the first-order autoregressive process, *Journal of Econometrics*, 13, 139-157.
- Goldberger, A.S. 1962, Best linear unbiased prediction in the generalised linear regression model, *Journal of the American Statistical Association*, 57, 369-375.

- Harvey, A.C., 1990, The econometric analysis of time series, LSE Handbooks in Economics (Phillip Allan, New York).
- Harvey, A.C., 1993, Time series models (Harvester Wheatsheaf, New York).
- Imhof, J.P., 1961, Computing the distribution of quadratic forms in normal variables, *Biometrika*, 48, 419-426.
- Koerts, J. and A.P.J. Abrahamse, 1969, On the theory and application of the general linear model (Rotterdam University Press, Rotterdam).
- Magnus, J.R. and B.Pesaran, 1991, The bias of forecasts from a first-order autoregression, *Econometric Theory*, 7, 222-235.
- Magnus, J.R. and T.J.Rothenburg, 1988, Least-squares autoregression with near-unit root, LSE typescript.
- Mandelbrot B.B., 1975, Limit theorems on the self-normalised range for weakly and strongly dependent processes, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31, 271-285.
- Mandelbrot B.B. and M.S.Taqqi, 1979, Robust R/S analysis of long run serial correlation, Invited paper, 42nd session of the International Institute of Statistics.
- Mandelbrot B.B. and J.W.Van Ness, 1968, Fractional Brownian motions, fractional noise and applications, *SIAM Review*, 10, 422-437.
- Messe, R.A. and K.J.Singleton, 1982, On unit roots and the empirical modelling of exchange rates, *Journal of Finance*, 37, 1029-1035.
- Nelson, C.R. and C.I.Ploser, 1982, Trends and random walks in macroeconomic time series: Some evidence and implications, *Journal of Monetary Economics*, 10, 139-162.
- Perron, P., 1988, Trends and random walks in macroeconomic time series : Further evidence from a new approach, *Journal of Economic Dynamics and Control*, 12, 297-332.

- Perron, P. 1991, Test consistency with varying sampling frequency, *Econometric Theory*, 7, 341-368.
- Shiller, R.J. and P.Perron, 1985, Testing the random walk hypothesis: Power versus frequency of observation, *Economics Letters*, 18, 1271-1289.
- Stock, J.H., 1994, Unit roots, structural breaks and trends, *Handbook of Econometrics*, Vol (4), ed. by R.F.Engle and D.Mcfadden, North Holland, New York, 2740-2841.
- Uhlig H., 1994, What macroeconomists should know about unit roots: A Bayesian perspective, *Econometric Theory*, 10, 633-644.

No.	Author(s)	Title
96105	H. Degryse	The Total Cost of Trading Belgian Shares: Brussels Versus London
96106	H. Huizinga and S.B. Nielsen	The Political Economy of Capital Income and Profit Taxation in a Small Open Economy
96107	T. Dieckmann	The Evolution of Conventions with Endogenous Interactions
96108	F. de Jong and M.W.M. Donders	Intraday Lead-Lag Relationships Between the Futures-, Options and Stock Market
96109	F. Verboven	Brand Rivalry, Market Segmentation, and the Pricing of Optional Engine Power on Automobiles
96110	D. Granot, H. Hamers and S. Tijs	Weakly Cyclic Graphs and Delivery Games
96111	P. Aghion, P. Bolton and S. Fries	Financial Restructuring in Transition Economies
96112	A. De Waegenaere, R. Kast and A. Laped	Non-linear Asset Valuation on Markets with Frictions
96113	R. van den Brink and P.H.M. Ruys	The Internal Organization of the Firm and its External Environment
96114	F. Palomino	Conflicting Trading Objectives and Market Efficiency
96115	E. van Damme and S. Hurkens	Endogenous Stackelberg Leadership
96116	E. Canton	Business Cycles in a Two-Sector Model of Endogenous Growth
9701	J.P.J.F. Scheepens	Collusion and Hierarchy in Banking
9702	H.G. Bloemen and E.G.F. Stancanelli	Individual Wealth, Reservation Wages and Transitions into Employment
9703	P.J.J. Herings and V.J. Vannetelbosch	Refinements of Rationalizability for Normal-Form Games
9704	F. de Jong, F.C. Drost and B.J.M. Werker	Exchange Rate Target Zones: A New Approach
9705	C. Fernández and M.F.J. Steel	On the Dangers of Modelling Through Continuous Distributions: A Bayesian Perspective
9706	M.A. Odijk, P.J. Zwaneveld, J.S. Hooghiemstra, L.G. Kroon and M. Salomon	Decision Support Systems Help Railed to Search for 'Win-Win' Solutions in Railway Network Design
9707	G. Bekaert, R.J. Hodrick and D.A. Marshall	The Implications of First-Order Risk Aversion for Asset Market Risk Premiums

No.	Author(s)	Title
9708	C. Fernández and M.F.J. Steel	Multivariate Student- <i>t</i> Regression Models: Pitfalls and Inference
9709	H. Huizinga and S.B. Nielsen	Privatization, Public Investment, and Capital Income Taxation
9710	S. Eijffinger, E. Schaling and M. Hoeberichts	Central Bank Independence: a Sensitivity Analysis
9711	H. Uhlig	Capital Income Taxation and the Sustainability of Permanent Primary Deficits
9712	M. Dufvenberg and W. Güth	Indirect Evolution Versus Strategic Delegation: A Comparison of Two Approaches to Explaining Economic Institutions
9713	H. Uhlig	Long Term Debt and the Political Support for a Monetary Union
9714	E. Charlier, B. Melenberg and A. van Soest	An Analysis of Housing Expenditure Using Semiparametric Models and Panel Data
9715	E. Charlier, B. Melenberg and A. van Soest	An Analysis of Housing Expenditure Using Semiparametric Cross-Section Models
9716	J.P. Choi and S.-S. Yi	Vertical Foreclosure with the Choice of Input Specifications
9717	J.P. Choi	Patent Litigation as an Information Transmission Mechanism
9718	H.Degryse and A. Irmén	Attribute Dependence and the Provision of Quality
9719	A. Possajennikov	An Analysis of a Simple Reinforcing Dynamics: Learning to Play an "Egalitarian" Equilibrium
9720	J. Jansen	Regulating Complementary Input Supply: Cost Correlation and Limited Liability
9721	J. ter Horst and M. Verbeek	Estimating Short-Run Persistence in Mutual Fund Performance
9722	G. Bekaert and S.F. Gray	Target Zones and Exchange Rates: An Empirical Investigation
9723	M. Slikker and A. van den Nouweland	A One-Stage Model of Link Formation and Payoff Division
9724	T. ten Raa	Club Efficiency and Lindahl Equilibrium
9725	R. Euwals, B. Melenberg and A. van Soest	Testing the Predictive Value of Subjective Labour Supply Data
9726	C. Fershtman and U. Gneezy	Strategic Delegation: An Experiment
9727	J. Potters, R. Sloof and F. van Winden	Campaign Expenditures, Contributions and Direct Endorsements: The Strategic Use of Information and Money to Influence Voter Behavior
9728	F.H. Page, Jr.	Existence of Optimal Auctions in General Environments
9729	M. Berliant and F.H. Page, Jr.	Optimal Budget Balancing Income Tax Mechanisms and the

No.	Author(s)	Title
		Provision of Public Goods
9730	S.C.W. Eijffinger and Willem H. Verhagen	The Advantage of Hiding Both Hands: Foreign Exchange Intervention, Ambiguity and Private Information
9731	A. Ridder, E. van der Laan and M. Salomon	How Larger Demand Variability may Lead to Lower Costs in the Newsvendor Problem
9732	K. Kultti	A Model of Random Matching and Price Formation
9733	J. Ashayeri, R. Heuts and B. Tammel	Applications of P-Median Techniques to Facilities Design Problems: an Improved Heuristic
9734	M. Dufwenberg, H. Norde, H. Reijnierse, and S. Tijs	The Consistency Principle for Set-valued Solutions and a New Direction for the Theory of Equilibrium Refinements
9735	P.P. Wakker, R.H. Thaler and A. Tversky	Probabilistic Insurance
9736	T. Offerman and J. Sonnemans	What's Causing Overreaction? An Experimental Investigation of Recency and the Hot Hand Effect
9737	R. Kabir	New Evidence on Price and Volatility Effects of Stock Option Introductions
9738	M. Das and B. Donkers	How Certain are Dutch Households about Future Income? An Empirical Analysis
9739	R.J.M. Alessie, A. Kapteyn and F. Klijn	Mandatory Pensions and Personal Savings in the Netherlands
9740	W. Güth	Ultimatum Proposals - How Do Decisions Emerge? -
9741	I. Woittiez and A. Kapteyn	Social Interactions and Habit Formation in a Model of Female Labour Supply
9742	E. Canton and H. Uhlig	Growth and the Cycle: Creative Destruction Versus Entrenchment
9743	T. Feenstra, P. Kort and A. de Zeeuw	Environmental Policy in an International Duopoly: An Analysis of Feedback Investment Strategies
9744	A. De Waegenaere and P. Wakker	Choquet Integrals with Respect to Non-Monotonic Set Functions
9745	M. Das, J. Dominitz and A. van Soest	Comparing Predictions and Outcomes: Theory and Application to Income Changes
9746	T. Aldershof, R. Alessie and A. Kapteyn	Female Labor Supply and the Demand for Housing
9747	S.C.W. Eijffinger, M. Hoeberichts and E. Schaling	Why Money Talks and Wealth Whispers: Monetary Uncertainty and Mystique

No.	Author(s)	Title
9748	W. Güth	Boundedly Rational Decision Emergence -A General Perspective and Some Selective Illustrations-
9749	M. Lettau	Comment on 'The Spirit of Capitalism and Stock-Market Prices' by G.S. Bakshi and Z. Chen (AER, 1996)
9750	M.O. Ravn and H. Uhlig	On Adjusting the HP-Filter for the Frequency of Observations
9751	Th. v.d. Klundert and S. Smulders	Catching-Up and Regulation in a Two-Sector Small Open Economy
9752	J.P.C. Kleijnen	Experimental Design for Sensitivity Analysis, Optimization, and Validation of Simulation Models
9753	A.B.T.M. van Schaik and H.L.F. de Groot	Productivity and Unemployment in a Two-Country Model with Endogenous Growth
9754	H.L.F. de Groot and R. Nahujs	Optimal Product Variety, Scale Effects, and Growth
9755	S. Hochguertel	Precautionary Motives and Portfolio Decisions
9756	K. Kultti	Price Formation by Bargaining and Posted Prices
9757	K. Kultti	Equivalence of Auctions and Posted Prices
9758	R. Kabir	The Value Relevance of Dutch Financial Statement Numbers for Stock Market Investors
9759	R.M.W.J. Beetsma and H. Uhlig	An Analysis of the "Stability Pact"
9760	M. Lettau and H. Uhlig	Preferences, Consumption Smoothing, and Risk Premia
9761	F. Janssen and T. de Kok	The Optimal Number of Suppliers in an (s,Q) Inventory System with Order Splitting
9762	F. Janssen and T. de Kok	The Fill Rate Service Measure in an (s,Q) Inventory System with Order Splitting
9763	E. Canton	Fiscal Policy in a Stochastic Model of Endogenous Growth
9764	R. Euwals	Hours Constraints within and between Jobs
9765	A. Blume	Fast Learning in Organizations
9766	A. Blume	Information Transmission and Preference Similarity
9767	B. van der Genugten	Canonical Partitions in the Restricted Linear Model
9768	W. Güth and B. Peleg	When Will the Fittest Survive? -An Indirect Evolutionary Analysis-
9769	E. Rebers, R. Beetsma and H. Peters	When to Fire Bad Managers: The Role of Collusion Between Management and Board of Directors

No.	Author(s)	Title
9770	B. Donkers and A. van Soest	Subjective Measures of Household Preferences and Financial Decisions
9771	K. Kultti	Scale Returns of a Random Matching Model
9772	H. Huizinga and S.B. Nielsen	A Welfare Comparison of International Tax Regimes with Cross-Ownership of Firms
9773	H. Huizinga and S.B. Nielsen	The Taxation of Interest in Europe: A Minimum Withholding Tax?
9774	E. Charlier	Equivalence Scales for the Former West Germany
9775	M. Berliant and T. ten Raa	Increasing Returns and Perfect Competition: The Role of Land
9776	A. Kalwij, R. Alessie and P. Fontein	Household Commodity Demand and Demographics in the Netherlands: a Microeconomic Analysis
9777	P.J.J. Herings	Two Simple Proofs of the Feasibility of the Linear Tracing Procedure
9778	G. Gürkan, A.Y. Özge and S.M. Robinson	Sample-Path Solutions for Simulation Optimization Problems and Stochastic Variational Inequalities
9779	S. Smulders	Should Environmental Standards be Tighter if Technological Change is Endogenous?
9780	B.J. Heijdra and L. Meijdam	Public Investment in a Small Open Economy
9781	E.G.F. Stancanelli	Do the Rich Stay Unemployed Longer? An Empirical Study for the UK
9782	J.C. Engwerda and R.C. Douven	Local Strong d -Monotonicity of the Kalai-Smorodinsky and Nash Bargaining Solution
9783	J.C. Engwerda	Computational Aspects of the Open-Loop Nash Equilibrium in Linear Quadratic Games
9784	J.C. Engwerda, B. van Aarle J.E.J. Plasmans	The (In)Finite Horizon Open-Loop Nash LQ-Game: An Application to EMU
9785	J. Osiewalski, G. Koop and M.F.J. Steel	A Stochastic Frontier Analysis of Output Level and Growth in Poland and Western Economies
9786	F. de Jong	Time-Series and Cross-Section Information in Affine Term Structure Models
9787	G. Gürkan, A.Y. Özge and S.M. Robinson	Sample-Path Solution of Stochastic Variational Inequalities
9788	A.N. Bancrjee	Sensitivity of Univariate AR(1) Time-Series Forecasts Near the Unit Root

P.O. BOX 00153 5000 LE TILBURG THE NETHERLANDS

Bibliotheek K. U. Brabant

17 000 01376876 8