THE SEASONAL UNIT ROOTS IN DEMOGRAPHIC TIME SERIES ANDTHE POSSIBILITIES OF ITS EXACT TESTING

Ondřej Šimpach, Petra Dotlačilová
University of Economics in Prague
ondrej.simpach@vse.cz, xdotp00@vse.cz

Key words:
seasonal unit roots – HEGY test – demographic time series – stationarity

Abstract:
The aim of this study is to present the options for testing seasonal unit roots in quarterly published demographic time series, in which the presence of seasonality is generally expected. The testing using sophisticated HEGY test will be presented in an econometric package GRETL and with selected demographic time series with quarterly frequency the absence of stationarity will be proved. Use of testing seasonal unit roots is necessary because the standard Dickey-Fuller test cannot be used in quarterly observed data in which the seasonality is expected. In the subsequent part of the study there will be compiled the model estimations for selected demographic time series.

Introduction

For the purposes of demographic analysis is needed to accept the assumption of non-stationarity in the demographic time series i.e. that in the time series the trend should occurred. The presence of the trend is one of the necessary requirements that analysed demographic time series can be used for modelling and eventual prediction. The presented article gives an exact opportunities to determine, whether the analysed demographic time series, published a quarterly frequency, signifies the stationarity or does not. There is not general rule, that every published quarterly time series is seasonal (see Arlt, Arltová [1]). The probability that the quarterly demographic time series is seasonal, however, is high, therefore, there will be presented approach of testing seasonal unit roots in quarterly published demographic time series, introduced by Hylleberg et al. [5]. The verification of stationarity is one of the conditions for further time series analysis, presented by Box and Jenkins [2] in methodological approach for time series modelling. In selected demographic time series are included the numbers of marriages, the numbers of divorces, numbers of live-born persons, the numbers of abortions, the numbers of deaths and the numbers of immigrants and emigrants. All observations were published by the Czech Statistical Office (CZSO). Analysed demographic time series start at first quarter 1991 and end at fourth quarter 2000 (i.e. 80 observations).

1. Methodology

At the sight of the trend of economic or demographic time series, we can hypothesize about the presence of stationarity of time series. If the trend is rising, respectively decreasing, and has its cycles, it is possible to say that a specific time series is non-stationary. Then also the autocorrelation function (ACF) has its first value very high, close to 1 and the other remaining values decrease very slowly. Graphical representation and definition of ACF is given e.g. by Arlt, Arltová [1].

In the case that the time series is non-seasonal, it is necessary exactly verified the presence of stationarity and it is possible to use the test of unit roots, presented by Dickey and Fuller in [3]. For the following models:

a) $X_t = \Phi X_{t-1} + a_t$
b) $X_t = c + \Phi X_{t-1} + a_t$
c) $X_t = c + \Upsilon_t + \Phi X_{t-1} + a_t$

where X_t is specific time series, c is constant and a_t is residue, the hypothesis are:

H_0: = 1 , i.e. the time series is I(1),
H_1: < 1 , is I(0).

Used test criterion

$$T = \frac{\hat{\Phi} - 1}{S_{\hat{\Phi}}} \tag{1}$$

has a *t*-distribution. However, due to the fact, that the tested hypothesis is "non-stationarity" the t-statistic has no standard *t*-distribution, but the distribution which was designed by Dickey and Fuller in [3]. In favour of the alternative hypothesis suggests low levels of test criterion *t*, respectively, values less than α-percent quantile of the distribution of D-F (see e.g. Granger [4]). In the case that the residual component in models a) b) or c) is auto-correlated then is constructed the extended DF test, which differs by extension by $\sum_{i=1}^{p-1} \Upsilon \, \Delta X_{t-i}$, so

a) $\Delta X_t = \Phi X_{t-1} + \sum_{i=1}^{p-1} \Upsilon \, \Delta X_{t-i} + a_t$
b) $\Delta X_t = c + \Phi X_{t-1} + \sum_{i=1}^{p-1} \Upsilon \, \Delta X_{t-i} + a_t$
c) $\Delta X_t = c + \Upsilon_t + \Phi X_{t-1} + \sum_{i=1}^{p-1} \Upsilon \, \Delta X_{t-i} + a_t$

and $\Delta X_{t-i} = X_{t-1} - X_{t-i-1}$ is the difference of neighbouring values.

In the case that we consider the seasonal time series, the Dickey-Fuller's methodological approach of unit root test cannot be used. Actually it is possible to use an approach, presented by Hylleberg et al. [5], (labelled HEGY), which was specially designed for testing for the presence of seasonal unit roots in quarterly observed time series Y_t. It is based on testing the statistical significance of the parameter π_i, where $i = 1, \dots, 4$ in regression equation, which according to Harvey and Dijk [6] may have the followed form:

$$\Delta_4 y_t = \mu_t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{3,t-1} + \sum_{j=1}^{k} \Phi_j \Delta_4 y_{t-j} + \varepsilon_t \qquad (2)$$

where $t = 1, \dots, T$. At the same time let us denote Δ_k, representing the filter, which can be defined as

$$\Delta_k y_t \equiv (1 - B^k) y_t \equiv y_t - y_{t-k} \qquad \forall\, k = 1,2, \dots \qquad (3)$$

where B is the lag operator. In the regression equation μ_t represents the deterministic trend. Now, let us denote

$$y_{1,t} = (1 + B + B^2 + B^3) y_t \qquad (4)$$
$$y_{2,t} = -(1 + B + B^2 + B^3) y_t \qquad (5)$$
$$y_{3,t} = -(1 - B^2) y_t \qquad (6)$$

and whereas the $(1 - B^4) = (1 - B) \cdot (1 + B) \cdot (1 + B^2)$, then y_t may contains seasonal unit roots. All filters, leading to $y_{1,t}$, $y_{2,t}$ and $y_{3,t}$ removing all unit roots except one which results from the fact, that the annual filter $(1 - B^4)$ can be decomposed as

$$(1 - B^4) = (1 + B + B^2 + B^3) \cdot (1 - B) \qquad \text{or}$$
$$(1 - B^4) = -(1 + B + B^2 + B^3) \cdot (1 + B) \qquad \text{or}$$
$$(1 - B^4) = -(1 - B^2) \cdot (1 + B^2)\,.$$

When in the regression equation (2) the parameter

- $\pi_1 = 0$, than the equation y_t contains non-seasonal unit root,
- $\pi_2 = 0$, than the equation y_t contains seasonal unit root in semi-annual frequencies, diminished by 1,
- $\pi_3 = \pi_4 = 0$, than the equation y_t contains seasonal unit roots in annual frequencies $\pm\, i$, where $i = 1, \dots, 4$.

Authors Hylleberg et al. [5] recommend the use of classical *t*-test to determine the statistical significance of the parameters π_1 and π_2 and next one *F*-test for joint statistical significance of the parameters π_3 and π_4. In econometric system Gretl can be used the HEGY add for testing the unit roots in seasonal time series. The tested hypothesis

H_0: the time series is non-stationary
H_1: non H_0

is expressed as

H_0: parameter z_1 ; z_2 ; z_3 ; z_4 $\neq 0$
H_1: non H_0

Compared to approach of authors Dickey and Fuller, where the null hypothesis is non-stationarity, in this case the tested hypothesis is stationarity. Gretl system has parameters π_1, π_2, π_3 and π_4 labeled as z_1, z_2, z_3 and z_4.

2. Testing of unit roots of analysed time series

From the selected demographic time series, which were published by CZSO with quarterly frequency (the numbers of marriages, the numbers of divorces, numbers of live-born persons, the numbers of abortions, the numbers of deaths and the numbers of immigrants and emigrants), were in all cases rejected the null hypothesis of stationarity at the 5% level of significance. At least one of the parameters z_1, z_2, z_3 and z_4 were at the 5% significance level equal to zero. The results are summarized in Table 1.

Given that in all analysed time series the non-stationarity was proved, the time series contain a trend so that can be modelled. The presented results were calculated in Gretl system, which is one of the few econometric systems, contains in addition the test for seasonal unit roots. Trend, which is included in the time series, can be seen with the naked eye from the images, presented at the end of this study (see Figure 1). Neither one of the analysed time series oscillate, except for marriages. This demographic time series has a tendency to fluctuate around the unconditional mean and could be suspected from the fact that does not have the trend. Exact test, however, confirmed that the trend contains and is non-stationary. The non-stationarity may be proved, among other things, by the fact that the amplitude increases over time.

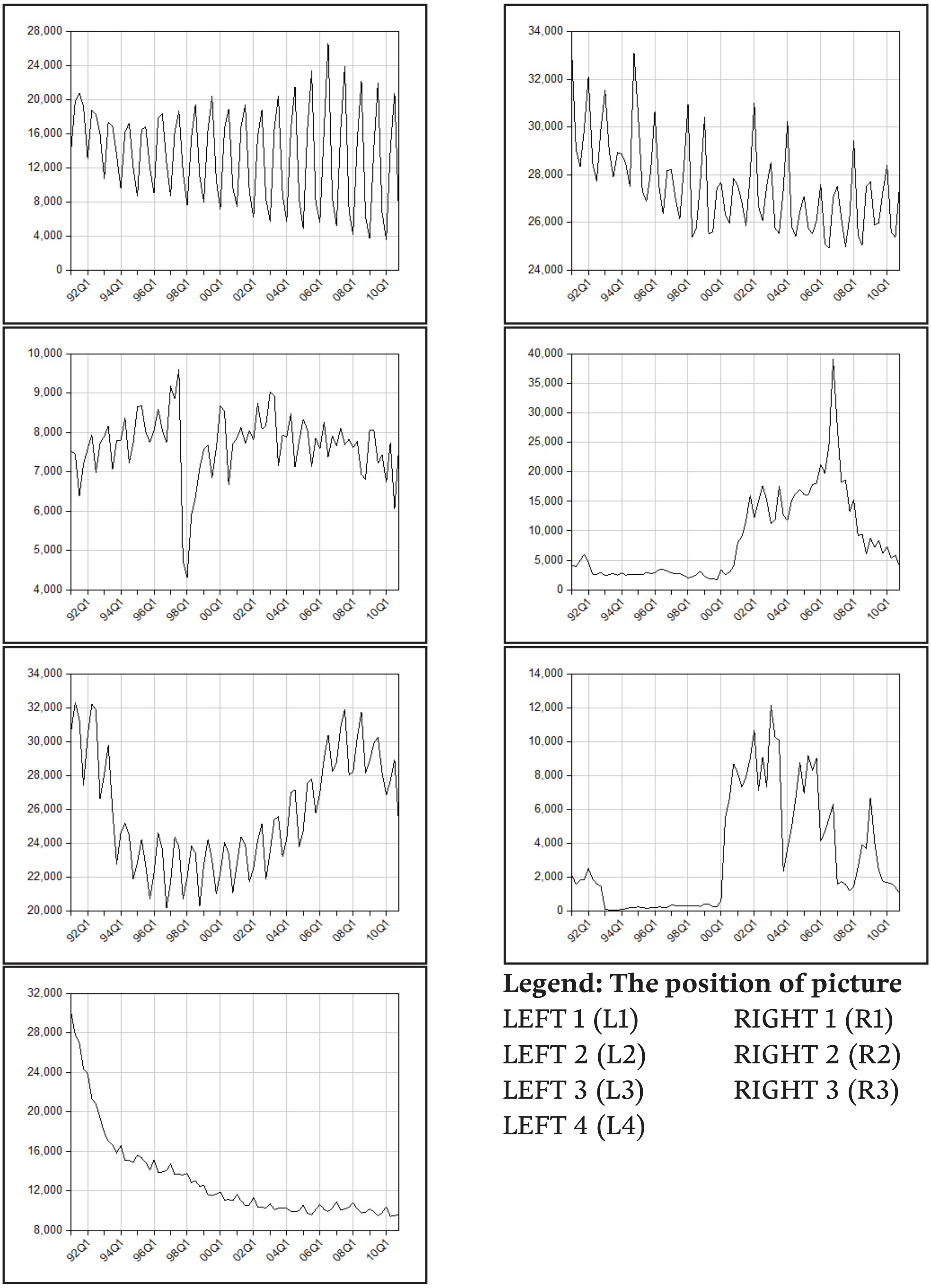
TAB. 1: **HEGY test results for selected demographic quarterly time series**

MARIAGES	**coefficient**	**standard error**	**t-stat.**	**p-value**
const	6345.35000	1841.35000	3.446	0.0010
z1	-0.11011200	0.02968890	-3.709	0.0004
z2	-0.02753800	0.03003090	-0.917	0.3624
z3	-0.00762274	0.00969591	-0.786	0.4345
z4	-0.01323320	0.00961451	-1.376	0.1732
d4y_1	0.42566300	0.09966930	4.271	0.0001
DIVORCES				
const	5053.850	1419.7700	3.560	0.0007
z1	-0.160991	0.0455561	-3.534	0.0007
z2	-0.205903	0.0885868	-2.324	0.0231
z3	0.220502	0.0959667	2.298	0.0247
z4	0.325187	0.0923045	3.523	0.0008
d4y_1	0.190322	0.1228250	1.550	0.1259

LIVE-BORN				
const	2283.4700	952.52400	2.397	0.0193
z1	-0.0276901	0.0101182	-2.737	0.0079
z2	-0.1776990	0.0855798	-2.076	0.0416
z3	0.0333594	0.0391601	0.852	0.3973
z4	0.0416267	0.0388963	1.070	0.2883
d4y_1	0.6023090	0.0938876	6.415	0.0000
ABORTIONS				
const	2951.8100	710.68900	4.153	0.0001
z1	-0.0518291	0.0116865	-4.435	0.0000
z2	-0.1387010	0.0590383	-2.349	0.0217
z3	0.0190186	0.0888773	0.214	0.8312
z4	0.2785010	0.0824795	3.377	0.0012
d4y_1	0.3841340	0.0915528	4.196	0.0001
DEATH				
const	18675.0000	7336.9800	2.545	0.0132
z1	-0.1623660	0.0622092	-2.610	0.0111
z2	-0.2576140	0.0776817	-3.316	0.0015
z3	0.0212165	0.0507676	0.418	0.6773
z4	0.1215390	0.0496249	2.449	0.0169
d4y_1	-0.0316371	0.1187440	-0.266	0.7907
IMMIGRANTS				
const	302.05700	825.39200	0.366	0.7155
z1	-0.0302791	0.0187532	-1.615	0.1110
z2	-0.5104460	0.1236140	-4.129	0.0001
z3	0.2699530	0.1143450	2.361	0.0211
z4	0.4748740	0.1032190	4.601	0.0000
d4y_1	0.1024440	0.1251530	0.819	0.4159
EMIGRANTS				
const	188.09000	422.94100	0.445	0.6579
z1	-0.0257016	0.0176582	-1.456	0.1501
z2	-0.1948100	0.0829772	-2.348	0.0218
z3	0.6054780	0.1136490	5.328	0.0000
z4	0.3641730	0.1274180	2.858	0.0057
d4y_1	0.0978633	0.1227800	0.797	0.4282

Source: author's calculation

FIG. 1: **The numbers of marriages (L1), divorces (L2), live-born persons (L3), abortions (L4), deaths (R1), immigrants (R2) and emigrants (R3)**



Legend: The position of picture

LEFT 1 (L1)	RIGHT 1 (R1)
LEFT 2 (L2)	RIGHT 2 (R2)
LEFT 3 (L3)	RIGHT 3 (R3)
LEFT 4 (L4)	

Source: author's construction

Conclusion

This study presented the possibilities for testing seasonal unit roots in quarterly published demographic time series, in which the presence of seasonality is generally expected. The sophisticated HEGY test was calculated in an econometric package Gretl and on selected demographic time series with quarterly frequency the absence of stationarity was demonstrated. Testing of seasonal unit roots is not old thing, it is a relatively young discipline. Determination of stationarity or non-stationarity is essential for the further continuation of the right analysis of demographic time series.

References:

[1] ARLT, J., ARLTOVÁ, M. *„Ekonomické časové řady“*, Grada Publishing, 2007.

[2] BOX, G.E.P., JENKINS, G. *„Time series analysis: Forecasting and control“*, San Francisco, Holden-Day, 1970.

[3] DICKEY, D. A., FULLER, W. A. „Distribution of the Estimators for Autoregressive Time Series with a Unit Root“, *Journal of the American Statistical Association*, 74, 1979, p. 427–431.

[4] GRANGER, C. W. J. „Some Properties of Time Series Data and Their Use in Econometric Model Specification“, *Journal of Econometrics* 16, 1981, p. 121-130.

[5] HYLLEBERG, S., ENGLE, R.F., GRANGER, C.W.J., YOO, B.S. „Seasonal Integration and Cointegration“, *Journal of Econometrics* 44, 1990, p. 215-238.

[6] HARVEY, D.I., van DIJK, D. *„Sample size, lag order and critical values of seasonal unit root tests“*, Loughborough University – Institutional Repository, 2003.