Eva Rubliková

HOW DETRENDING SERIES COULD INFLUENCE THE LENGTH OF CYCLE[1]

Abstract: *The aim of the article is to show how misunderstanding of the type generating process of the series could influence the length of the cycle in macroeconomic data. The focus of analysis is the study of the growth and cyclical behaviour of GDP in Slovakia based on the assumption that time series is trend or difference stationary. The results confirm that two different ways of separating cyclical fluctuations from the long-term trend have effect on the length of cycle. We display what different results could be obtained if we assume that the trend is deterministic but in fact it is stochastic. Our analysis is focused on 76 observations of seasonally adjusted quarterly data of GDP (bil. EUR in current prices) in Slovakia, during the period Q1/1995 till Q4/2013, using the decomposition method assuming that the series has two components – a long-term trend and stationary process containing the cycle. The long-term trend will be described by the linear trend function assuming trend stationary generating process and by the Hodrick – Prescott filter assuming the difference stationary generating process. After the estimation of the long-term trend and detrending the series, the residual stationary process will be analysed for the cycle and estimated by the autoregressive model of order 2 or higher.*

Keywords*: linear trend, Hodrick – Prescott filter, cycle, autoregressive model, Dickey-Fuller test*

JEL: C 22

1 Trend and difference Stationary Time Series

The modelling of trends and cycles in empirical statistical analysis of economic time series has a long history. Very often the secular movements used to be described by moving averages (simple or weighted) or by the various analytical functions of time. Cycles used to be described by goniometric functions or by the autoregressive models. Time series with trend, cycle and randomness are usually statistically

[1] Submission is supported by VEGA č. 2/0160/13 Finance stability and sustainability of economic growth in Slovakia in conditions of global economy and by VEGA č.1 /0285/14 Regional modelling of the economic growth of EU countries with concentracion on spatial econometric methods.

analysed by the decomposition method. The method used in this paper is based on [5], p. 119-210.

1.1 Trend Stationary Process

For classical trend-cycle decomposition is typically that time series, y_t observed over the period $t = 1, 2, ..., T$ is decomposed additively into a trend μ_t, and a cyclical component ε_t, which are assumed to be statistically independent of each other, i.e.

$$y_t = \mu_t + \varepsilon_t,\ \mathrm{cov}(\mu_t, \varepsilon_{t\pm k}) = 0 \text{ for all } t, k \tag{1}$$

In the model, y_t is the logarithm of the series which are usually observed annually, or quarterly (monthly) but seasonally adjusted. The simplest model for μ_t is the linear trend, written in the form

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{2}$$

Because y_t is measured in logarithms, β_1 assumes annual or quarterly (monthly) constant growth. Regression coefficients could be estimated by ordinary least squares and the cyclical component is then obtained by residuals as

$$\hat{\varepsilon}_t = y_t - \hat{\mu}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 t. \tag{3}$$

If cycles are present in the data, we assume that departures of y from μ_t must be only temporary, assuming that ε_t is stationary. According to the definition, stationary series requires that the mean of ε_t must be constant, which may be taken as zero here, the variance must be constant and finite, and covariances between ε_t and $\varepsilon_{t\pm k}$ must depend only on the time shift k:

$$E(\varepsilon_t) = 0 \quad E(\varepsilon_t^2) = \sigma_\varepsilon^2 < \infty \quad \gamma_k\ \mathrm{cov}(\varepsilon_t\, \varepsilon_{t\pm k}) = 0 \text{ for all } t \text{ and } k \neq 0. \tag{4}$$

Because ε_t is stationary time series, the series y_t is called trend stationary series or series with linear deterministic trend burried in stationary noise.

1.2 Difference Stationary Process

The random walk process with drift written by the first differences

$$\Delta y_t = \beta + \varepsilon_t \tag{5}$$

is known as difference stationary process. Accumulating the changes Δy_t from an initial value y_0 yields

$$y_t = y_0 + \beta t + \sum_{t=1}^{t} \varepsilon_t \tag{6}$$

In comparing models (2) and (6) we receive fundamental differences. The intercept y_t is not constant, but depends on initial value. The error ε_t is not stationary, because its variances and autocovariances increase with time. When $y_0 = 0$ a series has stochastic trend. If $y_0 \neq 0$ the model (6) represents a linear deterministic trend buried in non-stationary noise.

The distinction between trend and difference stationary processes has important implications for macroeconomic modelling of both economic growth and business cycles. If an observed series is difference stationary, then its growth component must itself be a non-stationary stochastic process rather than the more generally assumed deterministic trend. Instead of attributing all variation in a series to changes in the cyclical component, the difference stationary model allows for contributions from variation in both components.

1.3 How to Distinguish the Trend Stationary Series from Difference Stationary Series

Since the properties of the two classes of models are very different, it is necessary correctly distinguish between them. The resolution is very important, because it is very easy to confuse difference stationary series and trend stationary series that seems to provide a good fit to the data, with a high R^2, small residual variance and significant coefficients of deterministic trend, but which generates spuriously long cycles in the detrended data.

To distinguish whether an observed time series y_t is trend stationary or difference stationary, we can use the sample autocorrelation function with the first autocorrelation coefficient close to 1 or more exactly by the Dickey –Fuller unit root test. [5] p. 57.

1.4 The Hodrick – Prescott Filter

Other popular technique for extracting a cyclical component is the Hodrick – Prescott filter describing the secular trend when the series is difference stationary. [5] p. 92-95. This technique was used in the work [2] for GDP of Slovakia and the USA with the conclusion, that the log-linear trend and H–P filter used to give similar cyclical component, but the length of the cycle was not mentioned.

The Hodrick – Prescott filter was originally developed as the solution to the problem of minimising the variation in the cyclical component of an observed time series, $\varepsilon_t = y_t - \mu_t$, subject to a condition on the "smoothness" of the trend component, μ_t. This smoothness conditions penalise acceleration in the trend, so that the minimisation problem becomes that of minimising

$$\sum_{t=1}^{T} \varepsilon_t^2 + \lambda \sum_{t=1}^{T} [(\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1})]^2 \qquad (7)$$

with respect to μ_t, t = 0, 1, 2, ..., T+1, and where λ is a Lagrangean multiplier that can be interpreted as a smoothness parameter. The higher the value λ, the smoother is the trend. If $\lambda \to \infty$, μ_t becomes a linear trend. For quarterly data setting the smoothing parameter to $\lambda = 1600$.

We will denote the Hodrick – Prescott filter as H–P(λ). The finite sample H–P(λ) filter can be written as

$$y_t = \lambda\mu_{t-2} - 4\lambda\mu_{t-1} + (1+6\lambda)\mu_t - 4\lambda\mu_{t+1} + \lambda\mu_{t+2} \quad (8)$$

so that cycle is given as

$$\varepsilon_t = y_t - \mu_t = \lambda(\mu_{t-2} - 4\mu_{t-1} + 6\mu_t - 4\mu_{t+1} + \mu_{t+2} \quad (9)$$

This expression has to be modified for time series of the length t = 1,2, ..., T . Look to [5], p. 94-95.

The choice of λ = 1600 produces a filter that is close to the optimal for passing cyclical components having periods of 32 quarters or less, which corresponds to conventional views of the business cycle.

When filtering quarterly data with a near unit root, the optimal value is $\lambda \in$ (1000 - 1050), but the standard value of λ = 1600 does not lead to serious distortions.

2 Autoregressive Model of Cycle

Many realistic models for the cyclical component hidden in ε_t result from particular choices of the nonseasonal order p of the autoregressive model AR(p). Using back shift operator $B^p\varepsilon_t = \varepsilon_{t-p}$ we can express autoregressive model of order p in the form

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \ldots + \phi_p\varepsilon_{t-p} + u_t$$

or as

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)\varepsilon_t = (1 - g_1 B)(1 - g_2 B)\ldots.(1 - g_p B)\varepsilon_t = u_t \quad (10)$$

The roots $g_1, g_2, \ldots, g_p$ of the associated characteristic equation $\phi_p(B) = 0$ are pairs of complex numbers, which produce an autocorrelation function (ACF) following a damped sine wave and hence an ε_t containing cyclical fluctuations. Complex roots take the form $d\exp(\pm\pi fi$, and the autocorrelation function (ACF) becomes the damped sine wave

$$\rho_k = \frac{(sign(\phi_1))^k d^k \sin(2\pi fk + F)}{\sin F} \quad (11)$$

with damping factor $d = \sqrt{-\phi_2}$. The frequency f and the phase F of the wave could be obtained by the form

$$f = \frac{\cos^{-1}(|\phi_1|/2d)}{2\pi} \quad (12)$$

and

$$F = \tan^{-1}(\frac{1+d^2}{1-d^2}\tan 2\pi f) \tag{13}$$

respectively. Period of the cycle is then defined as $\frac{1}{f}$.

Autoregressive models of order p marked as AR (p) will describe an cyclical fluctuations if the model AR(p) can be factorised as

$$\phi_p(B)\varepsilon_t(1-d\exp(2\pi fi)B)(1-d\exp(-2\pi fi)B)\prod_{j=3}^{p}(1-g_jB)\varepsilon_t = u_t \tag{14}$$

3 Application to the Quaterly Seasonally Adjusted Data of GDP in Slovakia

Classical techniques in which the trend is described by the linear or nonlinear functions of time, or by the technique of moving averages with various lengths, were applied in [3] for GDP of V4 countries. There was demonstrated that the selection of the method for detrending the series used to influence the length of cycle in GDP series, varying from 4 to 9 years, but without explanation, what is the reason for that. By the following application we would like to explain, that the reason for such differences are hidden in misspecification of the process that generates the data.

Our analysis is concerned with quarterly data of GDP (in current prices), seasonally adjusted, from the period Q1/1995 till Q4/2013. The data are taken from the Statistical Office of the SR, the database SLOVSTAT. Development of GDP seasonally adjusted (GDPsadj) in SR is pictured in Figure 1 together with linear trend estimation for logarthmic transformation of GDPsadj.

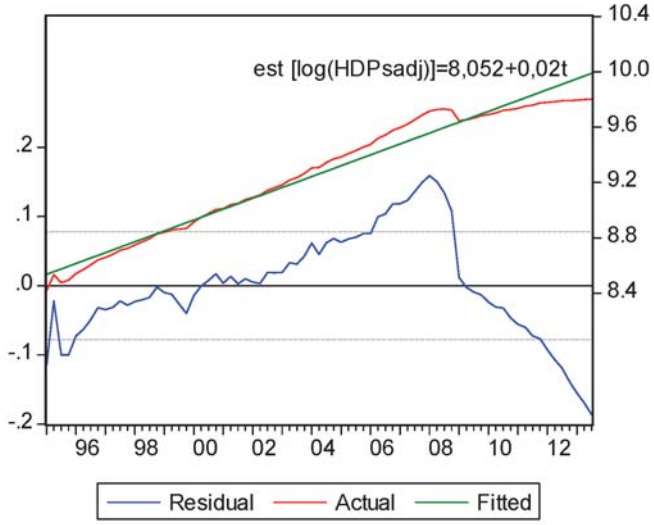
The log-linear trend gives statistically significant coefficients at any level of significance. The coefficient of determination $R^2 = 0{,}9682$ is close to unity, regardless of the actual rate of drift of the series or its variability. According to the results of estimated log-linear trend the mean annual growth is about 8%. This value is unrealistic, mainly during the years 2008–2013, when the world crisis started. The annual growth of GDP in 2009 comparing with the year 2008 was only 93.97 %, and during the years 2009 and 2013 the mean annual growth for GDP was 3.5 %. The residuals of the loglinear model are positively autocorrelated, with $DW = 0{,}08$.

Residuals pictured in Figure 1 show quite a long cycle and its AR(4) model has been estimated with results in Table 1.

The AR(4) model may be factorised as the pair of quadratics ($1 - 1.76B + 0.7994B^2$) and ($1 + 0.58B + 0.2605B^2$), which admit two pairs of complex roots. The first pair provides a cycle with a period 35.7 quarters (about 9 years), the second a period 6.5 quarters (about one and half years).

Figure 1

Series and the log-linear trend for GDPsadj



Source: Statistical Office of SR, database SLOVSTAT and own computations.

Table 1

Estimation of the cycle in residuals of log-linear trend of GDPsadj

Dependent Variable: CYCLEHDP
Method: Least Squares
Date: 03/14/14 Time: 17:43
Sample (adjusted): 1996Q1 2013Q3
Included observations: 71 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	1.171630	0.049278	23.77571	0.0000
AR(4)	-0.202890	0.052682	-3.851229	0.0003
R-squared	0.961036	Mean dependent var		0.004738
Adjusted R-squared	0.960471	S.D. dependent var		0.076449
S.E. of regression	0.015200	Akaike info criterion		-5.507336
Sum squared resid	0.015941	Schwarz criterion		-5.443599
Log likelihood	197.5104	Durbin-Watson stat		1.678077
Inverted AR Roots	.88-.05i	.88+.05i	-.29-.42i	-.29+.42i

Source: Own computation.

The estimated length of cycle is quite long, so the question is, whether the residuals are exactly stationary and whether the AR(4) model is appropriate.

Analysis of autocorrelation function for log-linear trend showed that the first coefficients of autocorrelation $r_1 = \hat{\phi}_{1,1} = 0.908 \rightarrow 1$, very close to 1. This is signal, that series of residuals has unit root and are not stationary. Also the first eight coefficients of autocorrelation function are statistically significant at 5% level of significance, given the limit $2/\sqrt{76} = 0.23$.

The first two partial coefficients of autocorrelation are statistically significant at 5 % level of significance, indicating AR(2) model, but this model did not give complex roots indicating the cycle.

The higher autoregresive model of order four shows the length of cycle about 9 years, which is quite long. The sum of coefficients AR(1) and AR(4) is 0.97 close to 1, indicating again that the residuals are nonstationary.

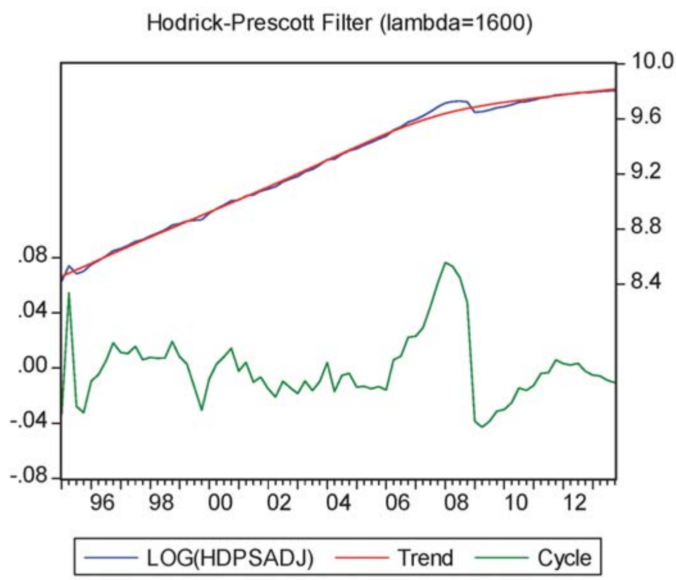
All mentioned results concerning the log-linear residuals indicate, that the GDP variable is generated by the random walk process with drift, or GDP variable has unit root, so it is stochastic variable and generated by the difference stationary process.

For that reason, we shall use Hodrick-Prescott filter, which alows for changing of trend and better estimation of the cycle.

Figure 2 shows the series GDPsadj with H-P(1600) filter estimated the long-term trend of the series and residuals (or cycle component of the series).

Figure 2

Detrending the series of log(GDPsadj) by the H-P(1600) filter



Source: Statistical Office of SR, database SLOVSTAT and own computations.

Detrended log(GDPsadj) series or residuals which express the cycle were analysed for stationarity using autocorrelation and partial autocorrelation function with the following results: the first autocorrelation coefficient 0.674 together with next two (0.519 and 0.329) are statistically significant. The first (0.674) and the fourth (-0.236) partial coefficient of autocorrelation are statistically significant too, because all of them in absolute value are larger than 0.23 taking empirical rule. The autoregressive model of order 4, have been estimated with the output in Table 2.

Table 2

Estimation of the cycle in detrended log(GDPsadj), by the Hodrick – Prescott filter

Dependent Variable: CYCLELGDPSADJ
Method: Least Squares
Date: 06/30/14 Time: 21:21
Sample (adjusted): 1996Q1 2013Q4
Included observations: 72 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.890784	0.066346	13.42634	0.0000
AR(4)	-0.199225	0.062825	-3.171099	0.0023
R-squared	0.723041	Mean dependent var		0.000545
Adjusted R-squared	0.719084	S.D. dependent var		0.023860
S.E. of regression	0.012646	Akaike info criterion		-5.875505
Sum squared resid	0.011195	Schwarz criterion		-5.812264
Log likelihood	213.5182	Durbin-Watson stat		1.780133
Inverted AR Roots	.77+.32i	.77-.32i	-.32+.43i	-.32-.43i

Source: Own computation.

Parameters of the model are statistically significant, with standard error of estimate S.E. = 0.012646, $R^2 = 0.72$ and DW = 1.78. Model AR(4) could be factorised as

$$(1 - 1.6B + 0.7556B^2)(1 + 0.66B + 0.3114B^2)$$

giving the length of two cycles: the first one is of the length 15.6 quarters or about four years and the second one is shorter, given 6.9 quarters or about 2 years.

This length of cycle seems to be realistic. Because the detrended series or residuals are stationary, the obtained results would be considered as results of model which fulfil our assumptions.

Conclusion

The time series of GDP seasonally adjusted in the SR during the period Q1/1985 till Q4 2013 were described by the logarithmic linear trend, and residuals series were described by the autoregressive model of order 4. We have found out that there are 2 cycles of the lengths 9 years and one and half years, but these results overestimate the right cycle, because the log-linear residuals are not stationary.

Using Hodrick – Prescott filter to obtain residuals in logarithmic transformation of seasonally adjusted GDP (LGDPsadj), we have found out again AR(4) model giving two cycles: one of the length about 4 years and the second one with the length of 2 years. The residuals were stationary, so this method gives better and realistic results.

We can conclude that it is very important to distinguish between trend and difference stationary processes. As we could see by this empirical analysis of cycles in GDP seasonally adjusted series, based on residuals from fitted trend lines, the results are confound with the two sources of variation, hence in general, used to overstate the magnitude and duration of the cyclical component and understate the importance of the growth component.

References

[1] HARVEY, A. C.: Trends and cycles in macroeconomic time series. In: *Journal of Business and Economic Statistics 3* (1985), 216-227.

[2] LUKÁČIK, M., SZOMOLÁNYI, K.: Možnosti analýzy hospodárskych cyklov. (Possibilities of business cycle analysis). In: *Forum Statisticum Slovacum* 7 (2011), p. 148-153.

[3] LUKÁČIK, M. – SZOMOLÁNYI, K.: Modelovanie trendov a cyklov v krajinách V4. (Modelling of trends and cycles in V4 countries). In: *Nové trendy v ekonometrii a operačním výskumu*. (New trends in econometrics and operational research). Bratislava : Vydavateľstvo EKONÓM, 2011.

[4] MILLS, T. C.: *Time series techniques for economists*. New York : Cambridge University Press, 1990.

[5] MILLS, T. C.: *Modelling Trends and Cycles in Economic Time Series*. New York : Palgrave Macmillan, 2003.