WORKING PAPERS AND STUDIES

COPYRIGHT

Luxembourg: Office for Official Publications of the European Communities, 2003

ISBN 92-894-5343-5
ISSN 1725-4825
Cat. No. KS-AN-03-015-EN-N

2003 EDITION

Statistical Methods for Potential Output Estimation and Cycle Extraction

EUROPEAN COMMISSION

Europe Direct is a service to help you find answers to your questions about the European Union

New freephone number:

00 800 6 7 8 9 10 11

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://europa.eu.int).

Luxembourg: Office for Official Publications of the European Communities, 2003

ISBN 92-894-5343-5
ISSN 1725-4825

TABLE OF CONTENTS

EUROPEAN COMMISSION
EUROSTAT

Directorate A: Statistical information systems; research and data analysis; technical co-operation with Phare and Tacis countries
Unit A-6: Statistical indicators for Euro-zone business cycle analysis

Statistical Methods for Potential Output Estimation and Cycle Extraction

By
Dominique Ladiray
Gian Luigi Mazzi
Fabio Sartori[†]

1. Introduction

Potential output is defined as the maximum level of durably sustainable production, without tensions in the economy, and more precisely without acceleration of inflation[1]. The output gap is the difference between effective production and the considered level of potential production. The potential production level is conceived as a supply indicator, and the output gap represents the excess (or the insufficiency) of demand. In this form, a positive number for the gap indicates excess demand and a negative number indicates excess capacity. The output gap represents transitory movements from potential output.

The analysis of the output gap is, in many cases, considered as the starting point for studying business cycles. The diagnosis of the economy in the cycle consists in evaluating the potential level of sustainable production without inflationary tensions, and then to compare it with the observed level of production.

Generally, the use of potential output and the output gap has the double ambition to point out the position of the economy within the cycle and to evaluate medium and long term growth. To summarise, output gap is estimated following two main approaches:

- The evaluation of the differences between the level of *actual* output and that of *potential* output, at a particular point in time. In this sense, output gap analysis provides information about excess capacity in the economy and it gives information about medium and long term economic growth.

[†] The views expressed herein are solely those of the authors and do not necessarily reflect the views of the European Commission. The authors would like to thank Roberto Astolfi, Gregory Czerwinski and Gabriella Manganelli for their valuable comments and contributions.

1 Adams et Al. 1987; IMF.

- The analysis of *cyclical fluctuations* of economic activity around *output trend*. In this regard, output gap analysis is an instrument to evaluate the adequacy of economic policy measures.

In practice, one may distinguish three possible main purposes for the estimation of potential output and the analysis of output gap:

- It is useful to the fine-tuning of policies according to the adjustment of demand and the support of supply.

- It is necessary to anticipate the inflationary risks to modulate accordingly the monetary policy.

- It makes it possible to calculate the structural public balance (that which would be observed if production were at its potential level) for the control of budgetary policy, public revenues/expenditures and excessive deficits.

In spite of the apparent consensus on the definition and the importance of potential output and output gaps the calculation of these variables is problematic, since potential output is not directly observable and neither is the output gap.

The calculation of potential output and of output gap conditional on and are sensitive to the model specification, the method of estimation and time horizon. Therefore the calculation necessarily relies on various statistical and theoretical hypotheses.

1.1 Theoretical assessment

An evaluation of the potential output is meaningful only if it supported by a coherent theoretical representation of the economy. Generally, we can consider that the potential output corresponds to the ideal equilibrium position for all the output variables; this particular position corresponds to the so-called "steady state".

The importance of the steady state and its interpretation can differ considerably following alternative economic theories. A very simple representation of such diversities can come to the following stylised interpretations:

- In a "keynesian" world the steady state, and by consequence the potential output, is a sort of asymptotic condition and the economy is normally fluctuating around it. Therefore the analysis of output gap reflects business cycle fluctuations around the long-run equilibrium.

- In a "monetary" world, the economy is supposed to be constantly at the steady state (equilibrium), therefore at its potential output. Only shocks can produce fluctuations around it. Therefore long-lasting shocks determine the potential output and transitory shocks enter the output gap.

The traditional macroeconomic approach depicts the economy as being keynesian in the short run and neo-classical in the long-term[2]. An extreme view of the debate would state that the more an economist is keynesian, the more he gives importance to output gap

---

[2] Cesse G., H Delessy(1997), "Ecarts de PIB: une grande variété de méthodes et de diagnostics".

analysis. On the contrary, a monetary economist would consider the divergence between potential and effective production of less significance.

Potential output is often assimilated to the so-called "permanent" component of the economy" which is typically non-stationary (see Nelson and Plosser 1982). At the same time, the output gap is often assimilated to the so-called "transitory" component of the economy", typically stationary but not necessarily periodical. This transitory component normally is composed of two main unobserved components: the cyclical and the irregular components. It is anyway useful to observe that, at this stage of our discussion, we are not considering any seasonal movement in the economic variables. If we accept that the economic variables can also be affected by seasonal variations, our transitory component will consist of three main components: cycle, irregular and seasonal.

We can identify two different ways of viewing economic fluctuations: the "trend deviation" interpretation of changes in overall production, and "gap closing" view on cyclical phenomena (Chagny and Döpke, 2001).

In this sense, when referring to the medium-term approach, the concept of potential output tends to be assimilated to that of trend and output gap to the deviation from the trend. Alternatively, if we focus on the short-run, the evolution of the output gap corresponds to the economic fluctuations, mainly related to the cyclical component in the economy.

From the cyclical point of view, the output gap analysis can identify the phase of the business cycle (acceleration/deceleration). By contrast, from trend point of view the output gap identifies the trend deviations and the joint analysis of potential output and the output gap (cycle/trend components) allows to detect the points of recovery/recession.

In order to avoid confusion, it should be clearly pointed out that the concepts of business cycle and trend are purely statistical, while the potential output and output gap derive from the economic theory.

Even if the features of cycles can be really similar, the consequences in terms of the impact of stabilisation policies are significant.

In the case of the "trend deviation", stabilisation policies can only reduce the variability of the observed data around the trend, without any possibility of influencing the growth. In fact in this view growth is determined by other factors, sometimes assumed exogenous. This is the typical scenario of optimal stabilisation policies and their performance is measured in terms of a quadratic loss function computed on the variables defined as a deviation from an equilibrium path.

In the case of "gap closing", stabilisation policies have also an effect on the growth component since trend and cyclical fluctuations are not independent. In this view, one important topic is represented by the definition of which type of fluctuations should be attributed to potential output and to the output gap. An extreme case is represented by real business cycle theory[3], which assumes that all fluctuations should be attributed to potential output, so that the output gap is represented only by random factors.

---

[3]Boschen J. and L. Mills (1990), "Monetary Policy with a New View of Potential GNP".

In the determination of output gap a key aspect is the time horizon over which the potential output is defined. In the simplest representation, growth of potential output is endogenous and relatively regular in the lung-run, since the determinants are not fixed: the capacity of utilisation depending on the technical progress and the labour factor on demographic growth.

Contrary, in the short-run those factors are considered to be exogenous. The potential output is determined in relation with the optimal combination of factors: the maximum utilisation of each factor not inducing inflationary tensions (“normal” capacity of utilisation) and “natural” unemployment rate.

In the following paragraphs a short description of the main issues in determining the definition of potential output are illustrated, of course, only a few have been pointed out: the time horizon, the relation between potential growth and unemployment (Okun’s law) and the unemployment rate (NAIURU).

1.1.1. Time horizon

The time horizon is a crucial issue in the definition of potential output: beyond the methodological diversity, the same approach can achieve a different evaluation of potential output and therefore of the output gap, depending on the reference time horizon and the frequency of data.

In particular, the estimation methods of the potential output can be concerned with a more or less long time horizon, by postulating higher or lower variability of the potential output. Indeed, the longer the reference time horizon is, the less the production factors are affected by the cyclical fluctuations and much more by structural factors.

The accepted variability of the potential output is inversely related to the time horizon. The longer the time horizon is, the lower is the variability of potential production and, accordingly, the higher is the amplitude of the variations of the divergence of the output. The same consideration can be made for the frequency of data.

Thus, if we are concerned with short-term inflationary tensions, we will tend to give priority to approaches accepting higher variability of potential output. If, on the other hand, we are concerned with a long-term growth scenario we will be able to accept a less volatile potential output.

Estimates of the output gap at particular points in time can vary considerably across estimation methods. On the other hand, however, there is considerable similarity in the broad time profile of the various potential output estimates. Indeed, since the level of the potential output can differ considerably depending on different estimations, it is difficult to identify the absolute size of the output gap but it is possible to infer its relative size.

An important consequence of these considerations is *that potential output and output gap do not refer to a universal evaluation*. Each analytical purpose can call on a specific degree of variability of potential output, and for given degree of variability, alternative approaches can be used.

For example, the Central Banks, as their principal purpose is the prevention of inflationary tensions, have *a priori* a shorter horizon, which implies not very flexible production factors. Therefore, in the calculation of output gap, Central Banks tend to prefer the cyclical perspective. As well, when Central Banks deal with the control of the budgetary policies, they privilege a longer time horizon and trends with lower volatility.

In a similar way, the International Monetary Found (IMF) and the World Bank, are mainly involved in the growth of developing countries and by consequence they focus on long term growth and their evaluation of the potential output has a much lower level of volatility.

1.1.2. The Okun's law

The article that provided the foundation for the potential output concept was published by A. Okun[4] in 1962. Okun proposed a simple linear relation between the divergence of the unemployment rate at its natural level and the divergence of production from its potential level. Therefore there is a negative correlation between changes in the unemployment rate and changes in output growth.

The unemployment/output relationship was more precisely considered by Okun as relating relative deviations of output from its potential level to deviations of the unemployment rate from its "natural" level. Potential output is meant to be the answer to Okun's question:

> *"How much output can the economy produce under conditions of full employment?*".

Because "full employment" is defined as the state in which labour markets are neither tight nor slack, inflationary pressures are presumed to arise when output growth pushes above its normal level, which in turn is related to declines in the unemployment rate below its normal level.

A key aspect of this perspective is the implicit, but crucial, role of the potential output and full-employment concepts in determining whether a particular growth rate or unemployment rate is inherently "inflationary".

The connection between unemployment and output growth is often formally summarized by the statistical relationship known as the "Okun' s law"[5]:

> *High GDP growth eventually places excessive strain on a nation's resources. This strain can become particularly acute in labour markets, where it is manifested as low unemployment. The labour market tightness associated with this low unemployment ultimately leads to higher prices.*

1.1.3. Phillips curve and NAIRU

The role of the output gap in affecting wage inflation was pioneered by Phillips (1958). The "Phillips curve" established an empirical relationship between price variation and the

---

4 Okun (1962) "Potential GNP: its Methods and Signifiance".

[5] David Altig, Terry Fitzgerald "Okun's Law Revisited: Should We Worry about Low Unemployment?", and Peter Rupert, May 1997.

unemployment level. More precisely the price variation, in the medium term, is equal to the variation of wages, after deduction of the productivity earnings.

In the event of perfect indexing of salaries to prices, there is only one unemployment rate which ensures that salaries grow at the same rate as the trend of labour productivity. This particular level of unemployment rate without inflationary tensions is commonly referred to as the NAIRU (Non Accelerating Inflation Rate of Unemployment).

In empirical work, inflation is often characterised as a mark-up over unit labour cost and imported goods prices, with the mark-up varying over the business cycle (de Brouwer and Ericsson, 1995).

Wages growth, in turn, also appears to be sensitive to the state of the business cycle and the rate of economic growth. This indicates that the output gap contains valuable information about movements in price and wage inflation. From a policy perspective, however, the underlying trend or potential output component should be defined in terms of a non-accelerating (or decelerating) inflation rate.

The NAIRU is stationary by nature, but in the event of temporary disequilibrium, one can suppose the need to increase the unemployment rate temporarily. Thus one can define two NAIRU concepts[6]:

1) a long term, NAIRU which considers only trend productivity.

2) a medium term NAIRU, with the possibility to be higher than the long term rate for a certain period.

2. Approaches to the estimation of potential output and output gap

In empirical analysis the definition of potential output and the output gap is much less strict than in economic theory. There is a large variety of statistical approaches to estimate such unobserved variables. The diversity of theoretical views on potential output and the output gap leads to a wide range of methods for their estimation. A first general classification has been proposed in Chagny and Döpke (2001). They distinguish the following categories of methods:

1. ***Univariate non-structural approaches***: methods that are based on some statistical procedure rather than referring explicitly to an economic theory (Cogley 1997). The interest in non-structural methods is partly motivated by the fact that they require less information than theory-based methods. This might be of relevance for the Euro-zone since there is still a lack of data at the aggregate level. Moreover, the methods can be implemented to model any time series of interest. This allows for a discussion of the cyclical behaviour of all parts of the economy, i.e. different types of expenditures and different sectors. Non-structural measures might therefore be used for a discussion of stylized facts of the business cycle.

[6] Hervé Le Bihan, Henri Sterdyniak, Philippe Cour "La notion de croissance potentitelle a-t-elle un sens?", CEPII Economie Internationale, 1997.

2. ***Direct measures of the cycle from survey data***: the potential growth and the production gap could be calculated by using business survey data. In this case, in a short time horizon, production technology is considered to be fixed and inputs to be complementary. Though the concrete questions in the surveys differ across the Euro-zone, the European Commission (2000) provides harmonised time series of industrial capacity utilisation. Thus, potential output equals effective output plus the gap between the available capacities and a level coherent with the absence of tensions on the goods market.

3. ***Structural approaches:*** these approaches rely on a specific economic theory, which is assumed to be correct. One can distinguish two broad groups of structural methods: multivariate methods with theoretical assumptions in so-called structural VARs (SVARs) and methods based on an aggregate production function. The approaches based on SVARs allow for more robust and reliable estimates of the output gap, since most of the underlying theories treat trend and cycle independently. Approaches based on production functions try to make explicit the nature of constraints that limit output (for example labour, capital, global factor productivity). Therefore, they require an analysis of the nature and the transmission of the disequilibria.

4. ***Multivariate non-structural approaches***: these approaches are mainly based on multivariate time series techniques and can be viewed as an extension and improvement of the univariate non-structural ones. Statistical relationships among different variables do not necessarily imply the acceptance of an economic theory.

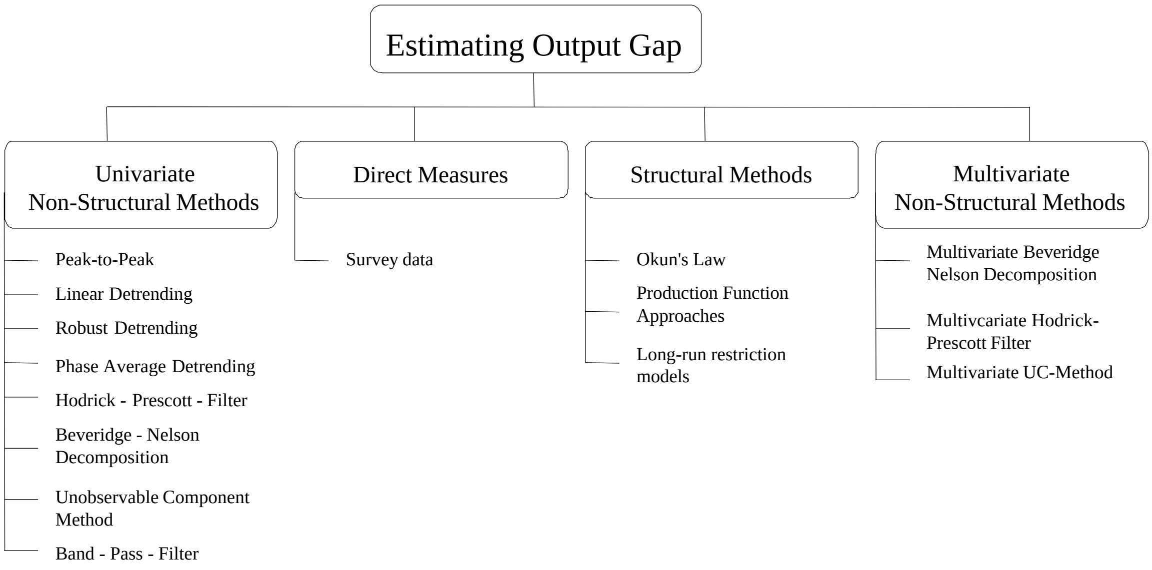

Figure 2.1 Approaches and methods for the estimation of the potential output

In this paper we concentrate our attention mainly on the first and last of the mentioned categories. Moreover we discuss long-run restrictions methods which are included in the third group. Concerning the remaining approaches proposed in the third category, they are outside the scope of this paper which is essentially focusing on statistical methodologies for estimating potential output and the output gap. Methods such as the Okun's law or those based on production functions require very strong economic priors.

The approach of using direct measures of the cycle from surveys is also outside the scope of this paper. This approach is based on a particular survey that measures the capacity of utilisation in industry. Essentially this survey defines potential output as the maximum allowable production level, which is not completely in line with the definition adopted in section 1. Moreover capacity utilisation data are only available for industrial sectors. The resulting output gap estimates appear quite volatile. Finally such surveys are based on subjective judgements from entrepreneurs and they cannot be considered as statistics in a strict sense. Nevertheless these surveys are an important source of information for economic analysis, as they provide a timely and reliable picture of the economic situation.

As shown in figure 2.1, several estimation methods for potential output have been developed within the different approaches. Several authors have contributed to this research topic starting from the end of the '70s.

We can classify these methods as follows:

1. Univariate methods

First difference filter	PAT filter, phase average trend (Boschan and Bry 1971)
Henderson filter	Stock and Watson decomposition (1986)
Hodrick and Prescott filter (1980)	Harvey decomposition (1985, 1989)
Baxter and King filter (1995)	Beveridge and Nelson decomposition (1981)

2. Multivariate methods

Hodrick and Prescott filter (Laxton and Tetlow 1992)	Beveridge and Nelson decomposition (1981)
SUTS decomposition (Harvey and Jaeger 1993)	Common trend and common cycle models (Vahid and Engle 1993, 1997)

This classification is based only on the number of time series involved in each method and not on their statistical properties. In other words this classification deals essentially with the amount of information used in the estimation of potential output and output gap. Clearly, univariate methods can be considered as "self explanatory" ones in the sense that they do not use any external and/or additional information. By contrast, multivariate methods use the information from many time series as well as relations derived from economic theory. Obviously all univariate methods are non-structural by definition, whereas multivariate methods can be split into structural and non-structural.

From a methodological point of view, we can introduce the following classification:

1. *Mechanical* methods based on filtering techniques;
2. *Mixed* methods based on filtering and time series techniques;
3. *Model based* methods, based only on time series techniques.

Method	**Type**	**Statistical properties**
Hoddrick and Prescott	Univariate/multivariate	Mechanical
Baxter and King	Univariate	Mechanical
PAT	Univariate	Mechanical
Detrending method	Univariate	Mechanical
Henderson method	Univariate	Mechanical
Stock and Watson decomposition	Univariate/multivariate	Mixed
SUTS method	Univariate/multivariate	Mixed
Beveridge and Nelson decomposition	Univariate/multivariate	Mixed
Common trend and common cycle decomposition	Multivariate	Model based

Table 2.1 Approaches and methods for the estimation of the potential output

3. Review of basic concepts concerning spectral analysis and linear filtering

In this section we review some concepts which are commonly used in the econometric analysis of time series, namely the population spectrum and linear filters. These concepts are part of a general approach to time series analysis called *analysis in the frequency domain* or *spectral analysis*, which is for many aspects dual to the more common approach, i.e. the *analysis in the time domain*.

Evaluating the problem of business cycle extraction from the spectral point of view will shed much light on the different methods presented in the following sections, will help understand their properties and will allow for a comparison of their performances. In fact we will introduce the crucial concept of *optimal filtering* in the frequency domain and we will see that many of the methods that have been proposed in the literature are simply an attempt to approximate an optimal filter.

3.1 Population spectrum and its main properties

Suppose y_t is a stationary process with mean μ and let γ_j denote the *j*th autocovariance of y_t such that:

(3.1) $$\gamma_j = E(y_t - \mu)(y_{t-j} - \mu)$$

If the autocovariances are absolutely summable, that is $\Sigma_{j=-\infty}^{\infty}\left|\gamma_j\right| < \infty$, then the following scalar function is well defined:

(3.2) $$\Gamma_Y(z) = \sum_{j=-\infty}^{\infty} \gamma_j z^j$$

where *z* is a complex scalar. $\Gamma_Y(z)$ is called the autocovariance generating function of y_t and, when divided by 2π and evaluated at $z = e^{-i\omega} = \cos(\omega) - i \cdot \sin(\omega)$ (where $i = \sqrt{-1}$ is the imaginary unit), gives the so called population spectrum of y_t:

(3.3) $$s_Y(\omega) = \frac{1}{2\pi}\Gamma_Y(e^{-i\omega}) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j e^{-i\omega j}$$

thus the population spectrum is a function of the real scalar ω; noticing that for any stationary process we have $\gamma_j = \gamma_{-j}$, simple trigonometric manipulations give a simplified expression for (3.3), that is:

(3.4) $$s_Y(\omega) = \frac{1}{2\pi}\left(\gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j \cos(\omega j)\right)$$

From (3.4) we see that the population spectrum is defined for any given value of ω; $s_Y(\omega)$ is a continuous, real-valued function of ω, periodic since $s_Y(\omega + 2k\pi) = s_Y(\omega)$ for any integer *k* and symmetric around $\omega = 0$. It is also possible to show that $s_Y(\omega)$ is nonnegative for any ω.

The jth autocovariance γ_j can be computed from the population spectrum using the following result:

(3.5) $$\gamma_j = \int_{-\pi}^{\pi} s_Y(\omega) e^{i\omega j} d\omega = \int_{-\pi}^{\pi} s_Y(\omega)\cos(\omega j) d\omega$$

Setting $j = 0$ in (3.5) leads to:

(3.6) $$\int_{-\pi}^{\pi} s_Y(\omega) d\omega = \gamma_0$$

that is, the area under $s_Y(\omega)$ between $\pm\pi$ is equal to γ_0, the variance of y_t. The result expressed by (3.6) can be generalised by calculating the integral of the population spectrum in a generic interval $\pm\omega^*$, with $0 \le \omega^* \le \pi$. In fact it can be shown that the quantity:

(3.7) $$\int_{-\omega^*}^{\omega^*} s_Y(\omega) d\omega = \int_{0}^{\omega^*} s_Y(\omega) d\omega$$

is equal to the portion of the variance of y_t that can be attributed to random fluctuations whose frequency lies in the interval $|\omega| \le \omega^*$.

3.2 Linear filters

One of the most important results of spectral analysis is that any stationary process y_t can be decomposed as an infinite sum of uncorrelated random components, each associated with a particular frequency. This result is known as *Cramér decomposition* and states that:

(3.8) $$y_t = \int_{-\pi}^{\pi} e^{i\omega j} dz(\omega)$$

where $dz(\omega)$ are complex orthogonal increments, that is they satisfy:

$$E(dz(\omega_1)\overline{dz}(\omega_2)) = 0 \text{ for every } \omega_1 \neq \omega_2$$
$$E(dz(\omega)\overline{dz}(\omega)) = s_Y(\omega) d\omega \text{ for every } \omega$$

where $s_Y(\omega)$ is the population spectrum defined in section 3.1.

Suppose that a second process x_t is obtained as a linear combination of past, present and future values of y_t, that is:

(3.9) $$x_t = \sum_{j=-\infty}^{\infty} h_j y_{t-j}$$

The relation (3.9) can be re-written introducing the lag polynomial $H(L)$ defined by:

(3.10) $$H(L) = \sum_{j=-\infty}^{\infty} h_j L^j$$

where L is the lag operator such that $L^k y_t = y_{t-k}$ for positive and negative values of k. Thus we obtain:

(3.11)
$$x_t = \sum_{j=-\infty}^{\infty} h_j L^j y_t = H(L) y_t$$

A lag polynomial like $H(L)$ in (3.10) is called a *linear filter* and x_t is the *filtered process* obtained from y_t by applying the filter $H(L)$. The Cramér decomposition for the filtered process x_t is given by:

(3.12)
$$x_t = \int_{-\pi}^{\pi} \tilde{H}(\omega) e^{i\omega j} dz(\omega)$$

where $\tilde{H}(\omega)$ is the Fourier transform or frequency response function of the linear filter $H(L)$, that is:

$$\tilde{H}(\omega) = \sum_{j=-\infty}^{\infty} h_j e^{-i\omega j}$$

For any given ω, $\tilde{H}(\omega)$ is a complex number that can be written in polar coordinates:

(3.13)
$$\tilde{H}(\omega) = G(\omega) e^{-i\Psi(\omega)}$$

In (3.13) the modulus $G(\omega)$ is called the *gain* of the filter $H(L)$ at frequency ω and measures the "amplification" induced by the filter on the components with frequency ω in the original series. On the other hand the *phase* $\Psi(\omega)$ measures the "time displacement" induced by the filter on the same component. Figure 3.1 shows the impact of gain and phase shifts over a specific periodic component of y_t.

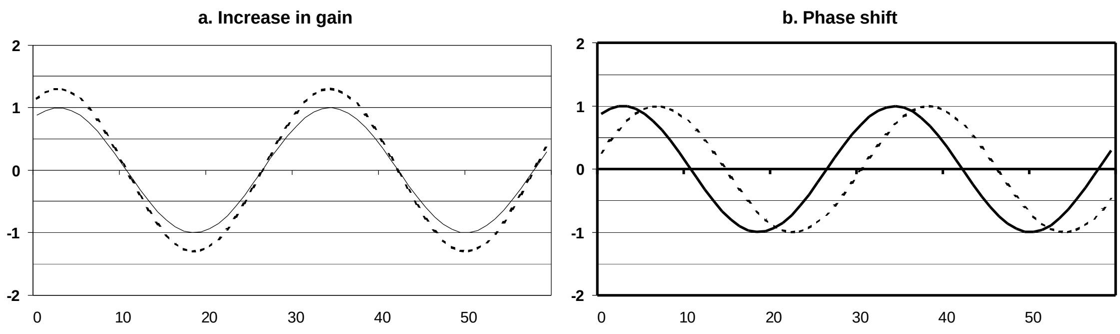

Figure 3.1 Effects of filtering (—original component, ······filtered component)

Let us consider the special case in which the filter defined (3.10) is *symmetric*, that is $h_j = h_{-j}$ for every j. It is easy to show that, for such a filter, the frequency response function $\tilde{H}(\omega)$ reduces to:

(3.14)
$$\tilde{H}(\omega) = h_0 + 2\sum_{j=1}^{\infty} h_j \cos(\omega j)$$

which is a real number for any ω. The important consequence is that a symmetric filter does not induce any phase shift at any frequency.

Suppose now that we want to construct a filter $H^*(L)$ that, when applied to a given series y_t, extracts only the components whose frequencies lie (in absolute value) in a specific subset Ω^* of $[0,\pi]$. Which are, from the spectral point of view, the conditions that must be imposed on $H^*(L)$ so that it performs this task in an optimal way? They are rather straightforward:

- the ideal filter is such that its gain is equal to one for the frequencies that must be extracted and zero elsewhere;

- the ideal filter is such that its phase shift is zero at every frequency.

The optimality conditions can thus be written as:

$$\begin{aligned} G^*(\omega) &= 1 \quad \text{for every } \omega \text{ such that } |\omega| \in \Omega^* \\ G^*(\omega) &= 0 \quad \text{for every } \omega \text{ such that } |\omega| \notin \Omega^* \\ \Psi^*(\omega) &= 0 \quad \text{for every } \omega \end{aligned} \tag{3.15}$$

According to the shape of Ω^*, a specific terminology is commonly used to indicate the ideal filters:

– if $\Omega^* = [0, \omega^*]$ with $0 \le \omega^* \le \pi$, then $H^*(L)$ fulfilling (3.15) is referred to as the *ideal low pass filter*, in the sense that it preserves the low frequency components of y_t and cuts off all high frequency components;

– if $\Omega^* = [\omega^*, \pi]$ with $0 \le \omega^* \le \pi$, then $H^*(L)$ is called the *ideal high pass filter*, which extracts from y_t only high frequency components and cuts off all low frequency ones;

– if $\Omega^* = [\omega_1^*, \omega_2^*]$ with $0 \le \omega_1^* \le \omega_2^* \le \pi$, then $H^*(L)$ is called the *ideal band pass filter*, which extracts from y_t only the components whose frequency lies in a specific range, cutting off all low and high frequency ones.

Figure 3.2 plots the gain function for the above mentioned ideal filters.

The cut-off frequencies for the ideal filters must be fixed according to the periodicity of the components we wish to extract and according to the frequency of the original data (annual, quarterly or monthly). For example if we want to extract from a quarterly series all components whose oscillation period lies between 1.5 and 8 years, then we need to fix $T_1 = 32\,(\text{period} = 8 \text{ years})$ and $T_2 = 6\,(\text{period} = 1.5 \text{ years})$, and the corresponding cut-off frequencies are $\omega_1^* = 2\pi/T_1 = \pi/16$ and $\omega_2^* = 2\pi/T_2 = \pi/3$.

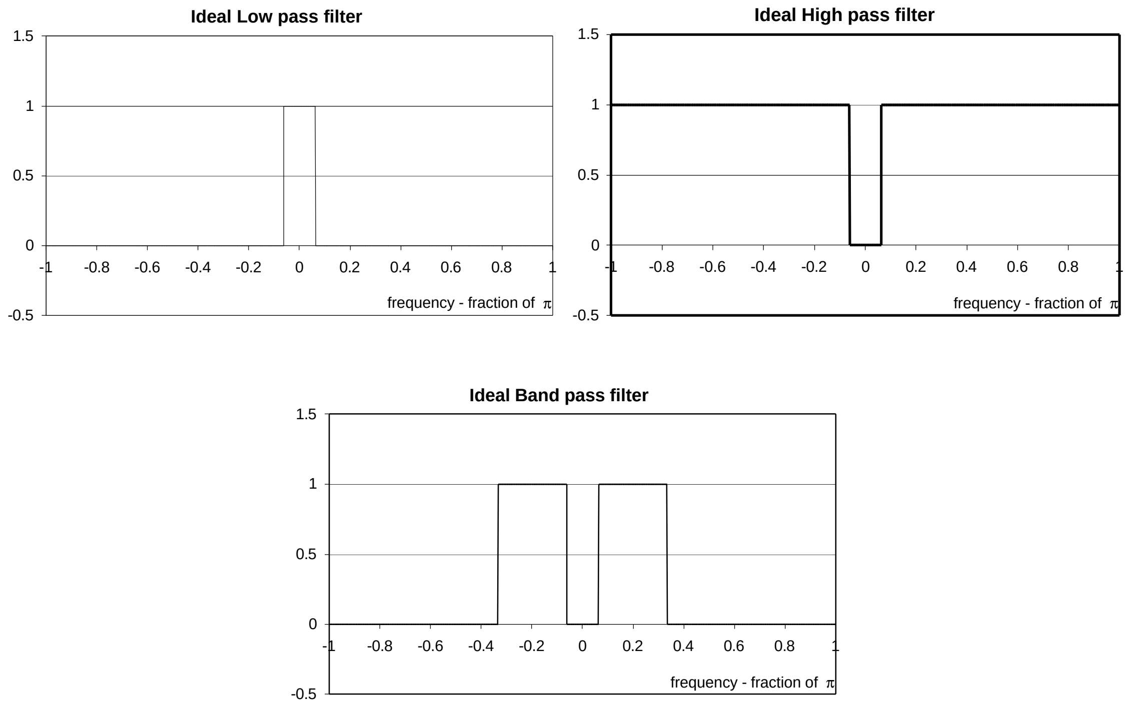

Figure 3.2 Gain of the ideal low pass, high pass and band pass filters

4. Univariate methods for the estimation of potential output and output gap

Letting y_t denote the series of interest, which is supposed to be seasonally adjusted and (eventually) log-transformed, we make the hypothesis that y_t can be decomposed, in an additive way, into a "permanent" component g_t, representing the growth or trend of y_t, and a residual or "transitory" component d_t, representing the deviation from the trend:

$$y_t = g_t + d_t$$

In a similar way the transitory component d_t can be decomposed in an additive way:

$$d_t = c_t + \varepsilon_t$$

where c_t is the cyclical component of the series and ε_t is the irregular component which is often modelled as white noise.

The decomposition for y_t is then:

(4.1) $$y_t = g_t + c_t + \varepsilon_t$$

In this section we will focus on the so-called univariate methods for cycle extraction. Some general remarks apply to these univariate methods:

- they only use information coming from observed output to infer the level of potential output;

- some of these methods, like for example the Hodrick and Prescott filter and the Beveridge and Nelson decomposition, explicitly consider the stochastic nature of potential output;

- these methods do not incorporate additional information about the state of potential output. In other words, they are subject to criticisms similar to those made on Okun's method in section 1.

4.1 Deterministic de-trending

The simplest hypothesis which can be formulated for the representation of the potential output is that it follows a deterministic (linear or quadratic) trend. In this case, all movements of the observed series y_t are attributed to the deviation component d_t since the output gap is represented by a completely deterministic function with fixed variations.

The practical relevance of such methods for the estimation of potential output is quite low. They are nevertheless briefly discussed in this paper since they can be viewed as a benchmark for more complex estimation methods.

In this context we assume that the potential output is a deterministic function of time. If the series y_t has been log-transformed and if we specify the trend as a linear function of *t*, that is:

(4.2) $$g_t = \beta_0 + \beta_1 t$$

then β_1 in (4.2) represents the average growth rate of the economy in the sample period.

It is important to observe that in this case we are able to obtain directly an estimation of the potential output g_t so that the gap component d_t is derived as a residual. This type of de-trending can be also viewed as a linear filter. It can be proved that the gain of this filter is characterised by the fact that it removes only partially the components whose frequency is close to zero.

4.2 First difference de-trending

Differencing data to eliminate the trend component from an economic time series is a quite common practice. If y_t is integrated of order one (or $y_t \sim I(1)$), that is $\Delta y_t = y_t - y_{t-1}$ is a stationary process, the application of the first difference filter removes the trend component from the series and gives an estimate of the transitory component. The transitory component d_t is estimated by:

(4.3) $$d_t = H(L)y_t = (1-L)y_t = y_t - y_{t-1}$$

The frequency response function of the first difference filter is:

$$\tilde{H}(\omega) = 1 - e^{i\omega}$$

and the corresponding gain is:

$$G(\omega) = \sqrt{2 - 2\cos(\omega)}$$

Figure 4.1 plots the gain of the first difference filter together with the gain of the ideal high pass filter. As is clear from the plot, the performance of the first difference filter as a high pass filter is quite poor when compared to the ideal filter. In fact the first differences over-emphasise the high frequency components of y_t, for which the gain is almost equal to two, and down-size considerably low frequency ones.

Moreover the filter is not symmetric, and its phase is nonzero. It can be shown that the phase shift induced by (4.3) is very high at low frequencies and decreases at higher ones.

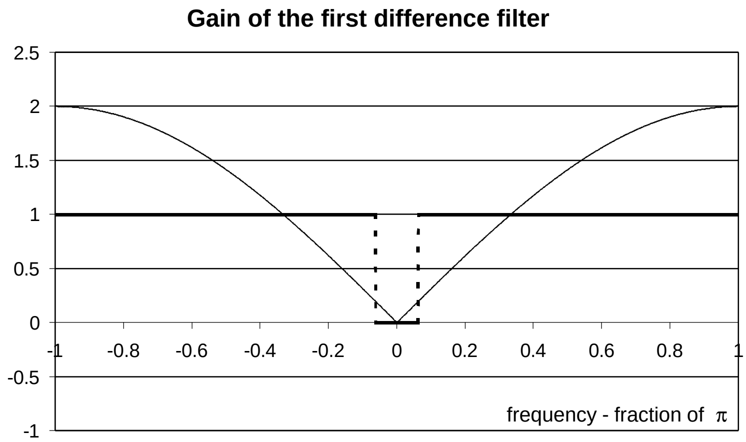

Figure 4.1 Gain of the first difference filter compared to the gain of the ideal high pass filter

De-trending using first differences gives just a rough idea of cyclical fluctuations. It is well known that when the data are nearly integrated, it can produce an over de-trending at zero frequency. This means that the auto-correlation function of the de-trended series is characterised by some negative spurious correlation at lag one. This situation can produce some relevant bias in the estimation of the cyclical component. In addition, if the data are stationary, the use of differentiation can produce spurious fluctuations which could mislead the economic interpretation.

4.3 Henderson moving averages

Henderson moving averages are used to extract the trend from an estimate of the seasonally adjusted series. They are based on a criterion which ensures a smooth estimation of the trend-cycle. Let us consider the series:

$$X_t = \begin{cases} 1 & \text{if} \quad t = 0 \\ 0 & \text{if} \quad t \neq 0 \end{cases}$$

Its transform by a centred moving average M of order $2p+1$ and with coefficients $\{\theta_i\}$, is given by:

$$MX_t = \begin{cases} 0 & \text{if } t < -p \\ \theta_t & \text{if } -p \le t \le p \\ 0 & \text{if } t > p \end{cases}.$$

This transform will therefore be smooth if the coefficients curve of the moving average is smooth.

Henderson (1916, 1924) proposed using quantity $H = \Sigma_i (\Delta^3 \theta_i)^2$ to measure the "flexibility" of the coefficients curve. This quantity is nil when the coefficients $\{\theta_i\}$ are located along a parabola and, in the general case, it measures the difference between the parabolic form and the form of the function giving the $\{\theta_i\}$.

Henderson then looked for order 2*p*+1 centred averages that retain the order-2 polynomials and minimise quantity *H*. The order 2*p*+1 Henderson moving average will be the solution of the minimisation problem:

$$\begin{cases} \underset{\theta}{Min} \Sigma_i (\Delta^3 \theta_i)^2 \\ \sum_{i=-p}^{i=+p} \theta_i = 1, \quad \sum_{i=-p}^{i=+p} i\theta_i = 0 \quad \text{and} \quad \sum_{i=-p}^{i=+p} i^2 \theta_i = 0 \end{cases}$$

The coefficients of these moving averages may also be calculated explicitly and, for an order 2*p*+1 average, by denoting $n = p + 2$, we have:

$$\theta_i = \frac{315\left[(n-1)^2 - i^2\right]\left[n^2 - i^2\right]\left[(n+1)^2 - i^2\right]\left[3n^2 - 16 - 11i^2\right]}{8n(n^2-1)(4n^2-1)(4n^2-9)(4n^2-25)}$$

Using this formula, it is therefore possible to calculate, in rational form, the coefficients of the Henderson moving averages. Therefore, for the sake of symmetry, presenting only the necessary coefficients, we have:

5 terms: $[5]\ ;\ \frac{1}{286}\{-21, 84, 160\}$

7 terms: $[7]\ ;\ \frac{1}{715}\{-42, 42, 210, 295\}$

9 terms: $[9]\ ;\ \frac{1}{2431}\{-99, -24, 288, 648, 805\}$

13 terms: $[13]\ ;\ \frac{1}{16796}\{-325, -468, 0, 1100, 2475, 3600, 4032\}$

And for 23 terms:

$$[23]\ ;\ \frac{1}{4032015}\{-17250, -44022, -63250, -58575, -19950, 54150, 156978, 275400, 392700, 491700, 557700, 580853\}$$

The coefficient curves, shown in figure 4.2, are smooth and the gain functions of these averages are closer to the ideal low-pass filter.

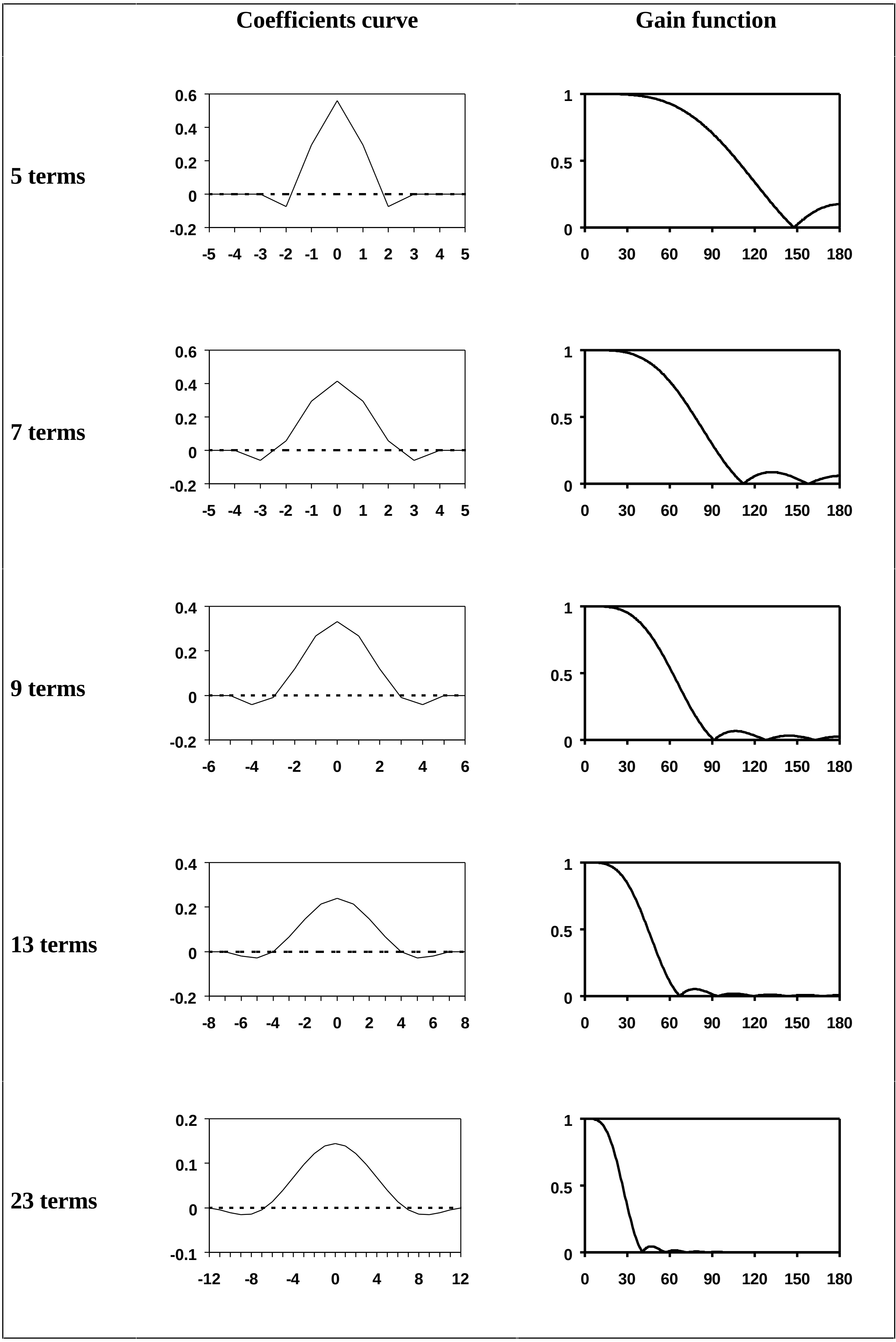

Figure 4.2 Coefficients and gain of the Henderson moving averages

4.4 The Phase Average Trend method

This method has been developed by the National Bureau of Economic Research (NBER) in the United States. It is extensively discussed in Boschan & Bry (1971) and Boschan and Ebanks (1978). It is extensively used also by the OECD to construct and analyse a system of cyclical indicators for the member countries, to identify and anticipate turning points. This method is essentially based on a recursive approach; furthermore it is the only method that provides the dating of turning points, thus giving an official chronology of the cyclical events.

4.4.1. Assessment of the Phase Average Trend Method

The Phase Average Trend Method (PAT) has three main objectives:

- the identification of cyclical turning points;
- the measure of long term trend;
- the construction of the so-called trend-adjusted series.

Let y_t be a seasonally adjusted series which can be decomposed like in (4.1). By applying an appropriate low-pass moving average, which will be presented below, it is possible to estimate the trend component g_t. The trend adjusted series d_t is equal to the transitory component:

$$d_t = y_t - g_t = c_t + \varepsilon_t$$

The identification of turning points is obtained by the comparison of different moving averages applied to an estimation of the transitory component d_t.

This comparison aims at avoiding that some artificial phase effects affect the results of the estimation: for example, the presence of irregularities in the original series may affect the dating of turning points. The identification of turning points gives the possibility to split the cyclical component in different phases, where the term "phase" refers here to the interval between two consecutive turning points of different sign and should not be confused with the same term introduced in section 3.2.

This splitting permits a new estimation of the permanent component g_t obtained by chaining up the average value of the series y_t of each phase. The Phase Average Trend method is essentially based on non-parametric techniques and it is mainly iterative. In its underlying philosophy this method is quite similar to the Census 2 seasonal adjustment procedure developed by the US Bureau of Census.

The OECD improved version of the Phase Average Trend method can be synthetically described by the following scheme.

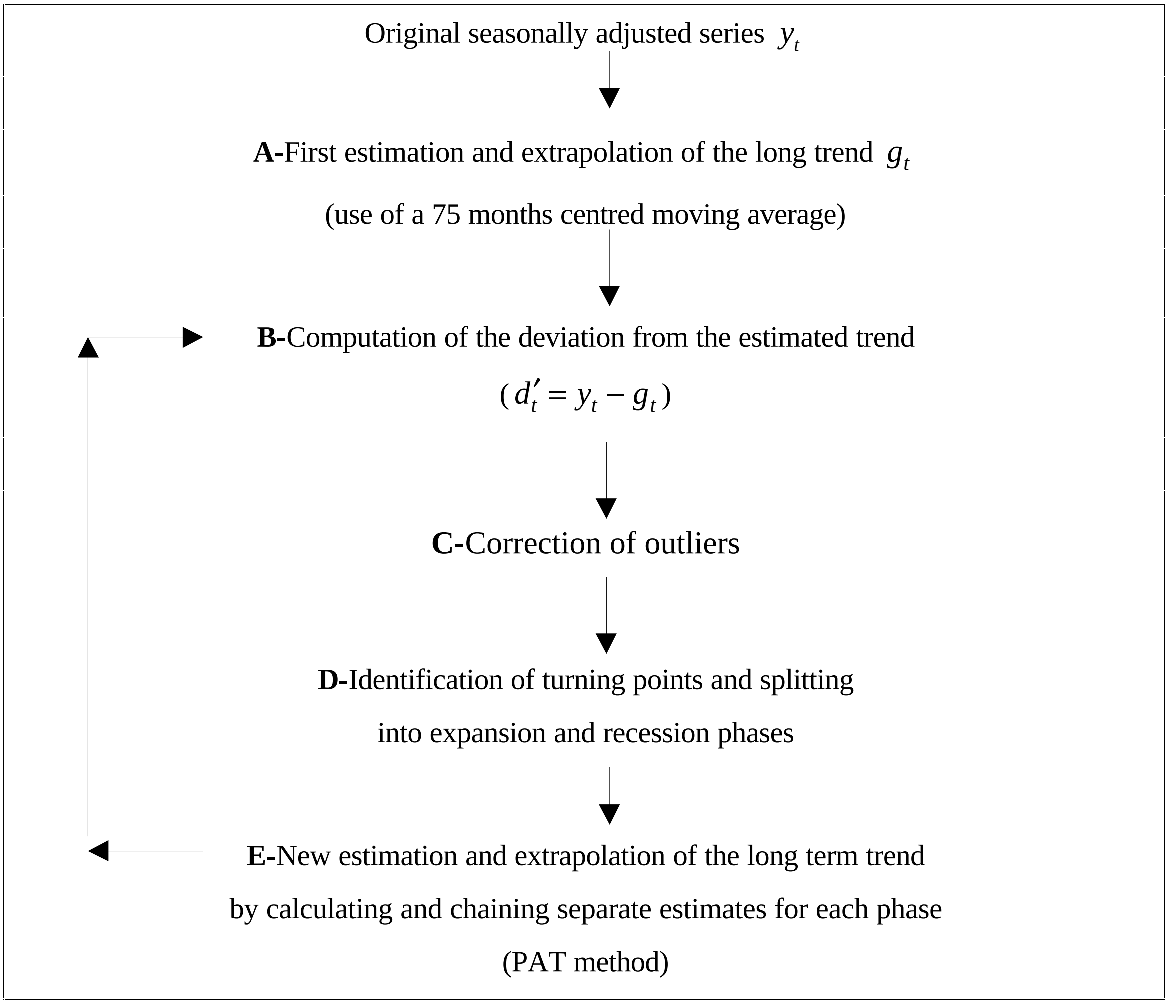

Figure 4.3 OECD improved version of the Phase Average Trend Method

The following points present some considerations on the different steps of Figure 4.3.

A-The long-term trend estimated via a 75 centred moving average with all the weight equal to 1/75:

$$g_t = \frac{1}{75} \sum_{j=-37}^{37} y_t L^j$$

The extrapolation is obtained using a quite rough method: more in detail the extrapolation is obtained by using projected monthly growth rates which appears absolutely inadequate. The choice of such moving average is essentially based on the assumptions made on the length of the business cycle fluctuation.

B-The component d_t is then estimated by $d'_t = y_t - g_t$.

C-The objective of this step is to obtain a first estimation of the cyclical component c'_t.

The sequence of the operations is as follows:

i) application to the estimated d_t' series of a so called Spencer 15 terms centred moving average:

(4.4) $$c_t' = \sum_{j=-7}^{7} h_j d_t' L^j$$

where the weights are given by:

h_j = [–3 , –6, –5, 3, 21, 46, 67, 74, 67, 46, 21, 3, –5, –6, –3] / 320

At the beginning and at the end of the series, the Spencer moving average is replaced by a truncated version. Spencer moving average is obtained by combining two moving averages of order 4 and 5 respectively. It gives the possibility to eliminate some infra-annual fluctuations.

ii) An indicator of regularity of the series d_t' is then computed. This indicator is called "months for cyclical dominance" (MCD). This indicator shows the number of months to be taken into account in order to ensure that the cyclical component is dominant on the irregular one. If T is the number of observations of y_t, we can compute over the whole period the mean of absolute growth rate. This mean can be computed for the cyclical component c_t' as obtained by (4.4), and to the estimated irregular component. The MCD indicator is then obtained by taking the first value of T for which the cyclical component is bigger than the irregular one. By convention it is assumed MCD $\leq$ 6.

iii) Given d_t' and c_t', an observation is considered an outlier if the value of each estimated irregular component is outside a range defined by its mean $\pm$ d times its standard deviation, where the scalar d depends from the volatility of the original series. Detected outliers are then replaced by the corresponding values of the Spencer estimation of the cyclical components c_t'. The resulting series is indicated by d_t''. This corrected series is again filtered by 15 terms centred Spencer moving average in order to obtain another estimation of the cyclical component c_t''.

iv) The series d_t'' is also filtered by using a 12 months moving average with all the weight equal to 1/12. This series will be defined as c_t'''.

At the beginning and at the end missing values are computed by simple truncated version of this filter. This filter is intended to eliminate any residual seasonality. In this way, we obtain two alternative evaluations of the cycle represented respectively by the application of the Spencer and of the 12 terms moving averages to the corrected series.

D-The turning point detection is obtained by comparing the turning points of the series obtained in step C. The idea is to identify turning points in a second moment on different smoothed time series in order to avoid the risk of any artificial displacement of the turning points. This method tries to avoid the risk of any wrong datation due to the presence of some irregularities. The sequence is as follows:

1) On the series c_t''' a first group of turning points, called potential turning points (PTP) is identified. A potential peak or slack is defined by the fact that it is higher (lower)

than any other point in the range of the five preceeding or following months. The sequence of peaks and slacks is checked.

2) The so-called corresponding turning points (CTP) are identified in the series c_t''. All CTP in the range of ± 5 months around a PTP are retained, all turning points identifying cycles of less than 15 months are eliminated. The sequence of peaks is checked again.

3) A new moving average of the same order of the MCD (less or equal thant six months) indicator is applied to the non-corrected series d_t'. A simple extrapolation is performed at the beginning and at the end of the series. The new series is called short-term moving average. On the resulting series, a new set of corresponding turning points (CTP1) is identified. The second part of step 2 is repeated comparing these CTP1 to the CTP (used as a benchmark).

4) A last version of corresponding turning points (CTP2) is derived directly from the series d_t. All turning points in a range of $\pm k$ months around the CTP1 are retained, where $k = 4$ if MCD≤4 and k = MCD if MCD>4.

The second part of step 2 is applied again to the CTP2 with the CTP1 as a benchmark. All turning points occurring during the first and the last 6 months are eliminated. All turning points creating a phase of less than 5 months or cycles of periodicity less than 15 months are also eliminated. This new chronology of turning points CTP2 is called tentative turning points (TTP).

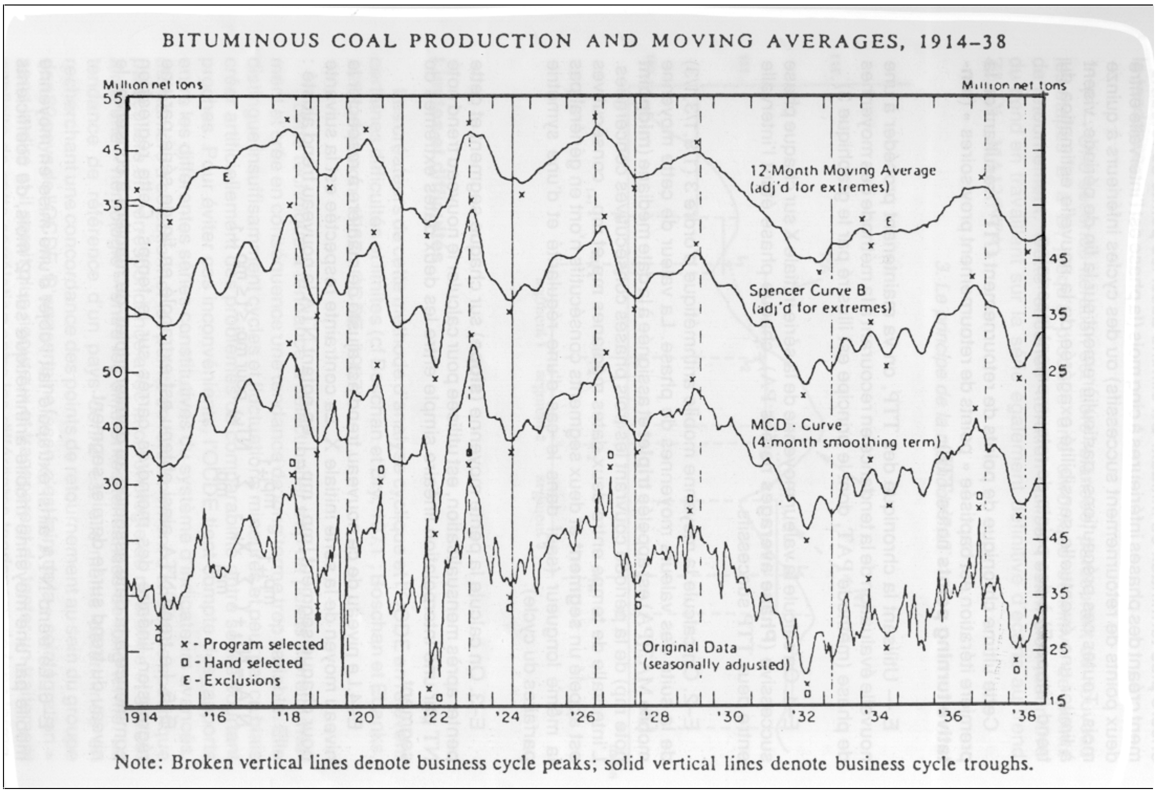

Figure 4.4 Identification of cyclical turning points of an economic time series

E- A new evaluation of the long term trend is obtained starting from the TTP first chronology. This step, which is probably the core of the Phase Average Trend method, essentially calculates an average slope for the series in each phase. This average slope is then used to obtain a new estimate of the trend component, with the constraint that its average must be equal to the average of the original series.

E1- At the first stage the average value of the original series y_t on each phase is computed. The resulting series is called PA (Phase Average).

E2- A simple three terms moving average with the weights equal 1/3 is computed on the average values of each phase. The value of this moving average is assigned to the medium value of the current period. The time interval between two consecutive medium points is called segment. Two consecutive segments do not have normally the same length (without the case in which the cyclical fluctuation are perfectly symmetric and regular).

E3- The slope is then computed in each segment. The slope is used to compute the New Trend called NT by a simple linear interpolation between the two extreme date of each segment.

E4- The level of this New Trend is then adjusted to respect the average level of the original series y_t. This adjusted trend is called New Trend Adjusted, NTA.

E5- The NTA is then extrapolated at the end of the series by using a linear regression on time. This regressions starts on the date associated to the last legal point in order to respect the level of the trend on the last segment.

E6- The extrapolated NTA is smoothed by a 12 terms moving average with weights equal to 1/12. The main objective is to smooth the link between two consecutive segments. At the beginning and at the end of the series is simply extrapolated. The resulting series constitute the final estimation of the trend.

Steps **B** to **D** are iterated to produce a final chronology of turning points. This iteration starts from the final estimation for the trend obtained in step **E** and uses the definitive estimation of the trend deviation.

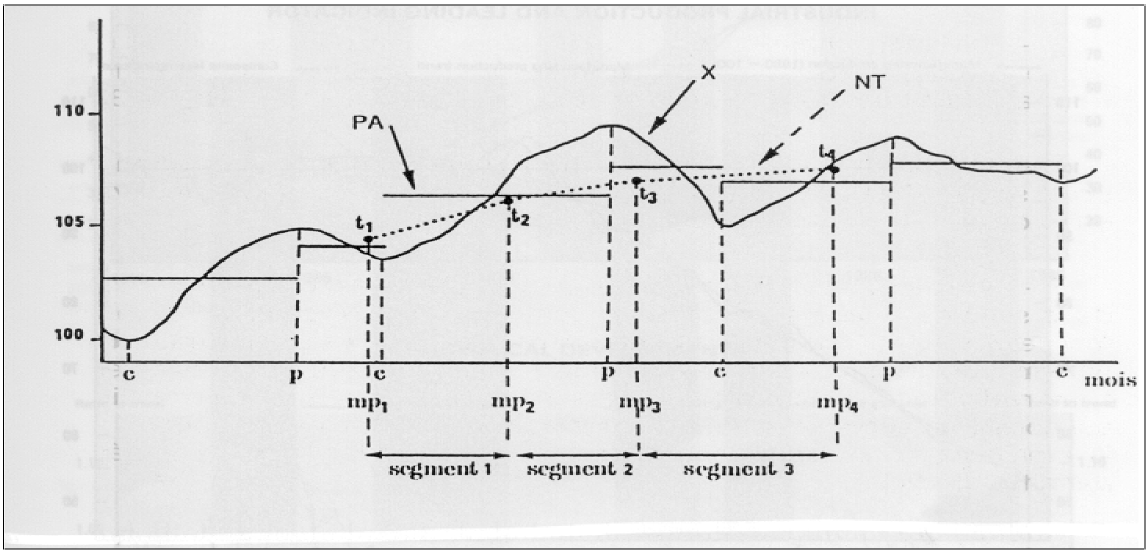

Figure 4.5 description of the Phase Average Trend method.

4.4.2. Some remarks on the Phase Average Trend methodology

We can point out some open problems with the PAT approach:

- the number of detected turning points is often too high and the estimated long term trend is often too unstable;
- the identification of new turning points can considerably affect the long term trend estimation. The long term trend estimation is too "fragile" in particular at the end of the period;
- the extrapolation of the trend can result considerably biased;
- on the other hand this method allows a quick identification of local inflexion of the trend, and this is more evident by analysing historical fluctuations;
- finally, this is the only method, among those presented in this review, that provides an official chronology of turning points.

4.5 Exponential smoothing

This de-trending method has been widely used in the past by economists: among others, Friedman (1957) applied it in the context of a permanent income analysis and Lucas (1980) in various empirical studies. An overview of this method, with an emphasis on its close relation to the Hodrick and Prescott filter presented in section 4.6, can be found in King and Rebelo (1993). The exponential smoothing filter gives the permanent component of the series y_t as the solution to the following minimisation problem:

(4.5) $$\min_{\{g_t\}_{t=1}^{T}} \sum_{t=1}^{T} \left[(y_t - g_t)^2 + \lambda (g_t - g_{t-1})^2 \right]$$

The parameter λ penalises the changes in the permanent component (or potential output) g_t. The degree of smoothness of the estimated g_t depends strongly on the chosen value for λ. The first order conditions for the minimisation problem (4.5) take then the form:

$$-2(y_t - g_t) + 2\lambda(g_t - g_{t-1}) - 2\lambda(g_{t+1} - g_t) = 0$$

Such conditions show the link between the transitory component $d_t = y_t - g_t$ and the changes in the permanent component g_t.

In order to study the characteristics of corresponding low pass and high pass filters the first order conditions can be rewritten in the form:

$$y_t = F(L) g_t$$

where $F(L)$ is given by:

$$F(L) = -\lambda L^{-1} + (1 + 2\lambda) - \lambda L = 1 + \lambda(1 - L)(1 - L^{-1})$$

The exponential smoothing approach can be viewed as a two-side filter able to render stationary series characterised by stochastic trends (Nelson and Plosser, 1982) or, more generally integrated series up to order two (see King and Rebelo 1993).

King and Rebelo (1993) proposed a detailed analysis of this filter in both frequency and time domain. In particular they computed the gain functions of the filter showing that it is not a good approximation of the optimal filter.

4.6 The Hodrick and Prescott filter

The Hodrick and Prescott filter is the most known and commonly used univariate method for the estimation of potential output. It is largely used in scientific papers as well as by international organisations like the IMF and the OECD. In the European Union it is used by the Economic and Financial Affairs Directorate and in the Economic Directorate of the European Central Bank.

4.6.1. Derivation of the Hodrick and Prescott filter

The application of the Hodrick and Prescott filter extracts from y_t the growth component g_t. The estimation of g_t is obtained through the minimisation of the sum of squares of the transitory component subject to a penalty for the variation in the second differences in the growth component. That is g_t is the solution to the following minimisation problem:

(4.6) $$\min_{\{g_t\}_{t=1}^{T}} \sum_{t=1}^{T} \left[(y_t - g_t)^2 + \lambda[(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 \right]$$

where λ is a penalty parameter which is closely related to the "smoothness" of the estimated trend. With our notation (4.6) can be rewritten as:

$$\min_{\{g_t\}_{t=1}^{T}} \sum_{t=1}^{T} \left[d_t^2 + \lambda(\Delta^2 g_{t+1})^2 \right]$$

The minimisation of (4.6) leads to a system of linear equations giving the series y_t as a function of its permanent component via a $T \times T$ matrix $\boldsymbol{M}$:

(4.7) $$\boldsymbol{y} = \boldsymbol{M}\boldsymbol{g}$$

Where $\boldsymbol{y}$ and $\boldsymbol{g}$ are, respectively, the series y_t and g_t stacked in a column vector. The first order conditions for the minimisation of (4.6) give:

$$\begin{aligned}
&d_1 = y_1 - g_1 = \lambda(g_1 - 2g_2 + g_3) \\
&d_2 = y_2 - g_2 = \lambda(-2g_1 + 5g_2 - 4g_3 + g_4) \\
&d_t = y_t - g_t = \lambda(g_{t-2} - 4g_{t-1} + 6g_t - 4g_{t+1} + g_{t+2}) \quad \text{for } t = 3,\ldots,T-2 \\
&d_{T-1} = y_{T-1} - g_{T-1} = \lambda(g_{T-3} - 4g_{T-2} + 5g_{T-1} - 2g_T) \\
&d_T = y_T - g_T = \lambda(g_{T-2} - 2g_{T-1} + g_T)
\end{aligned}$$

that is, in matrix form:

$$\boldsymbol{y} - \boldsymbol{g} = \lambda \boldsymbol{F}\boldsymbol{g}$$

where:

$$\underset{(T\times T)}{\boldsymbol{F}} = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & \dots & 0 \\ -2 & 5 & -4 & 1 & 0 & 0 & \dots & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & \dots & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

The permanent component is given by:

(4.8) $$\boldsymbol{g} = \boldsymbol{M}^{-1}\boldsymbol{y} = (\lambda\boldsymbol{F} + \boldsymbol{I})^{-1}\boldsymbol{y}$$

and the matrix ***M*** depends only on the number of observations and on the parameter λ.

It is easy to understand that for this filter the role of the smoothing parameter λ is crucial. By increasing the value of λ we obtain smoother estimates of the growth component g_t and more volatile estimates of the transitory component d_t. In fact if $\lambda \to \infty$, (4.6) is minimised if the estimated trend is a straight line (for which $\Delta^2 g_t$ are identically zero); on the other hand if $\lambda = 0$ then (4.6) is minimised if $g_t = y_t$ for every *t*. Figure 4.6 gives an example of three different estimates of the trend component corresponding to different choices for λ.

In their original paper Hodrick and Prescott (1980) propose some recommended values for λ. They suggest:

- λ = 100 for annual data;
- λ = 1600 for quarterly data;
- λ = 14400 for monthly data.

This means that the proposed values for λ are given by $100 \cdot f^2$ where *f* is the frequency of y_t (*f* = 1, 4 and 12 for annual, quarterly and monthly data). Further insights on the choice for λ will be given in the next section. Figure 4.6 shows the estimated trend component for the quarterly index of industrial production in the Euro-zone arising from three different choices for λ and figure 4.7 compares the three corresponding transitory components.

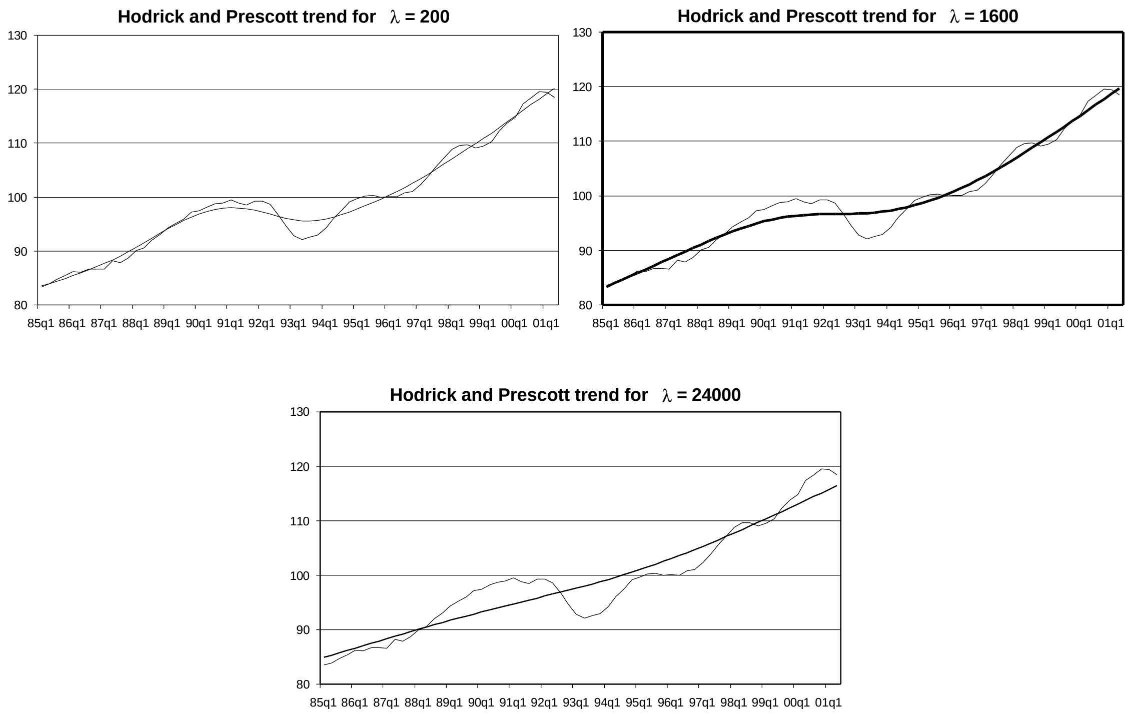

Figure 4.6 Estimated trend of the quarterly index of industrial production in the Euro-zone applying the Hodrick and Prescott filter with $\lambda = 200$, $\lambda = 1600$ and $\lambda = 24000$.

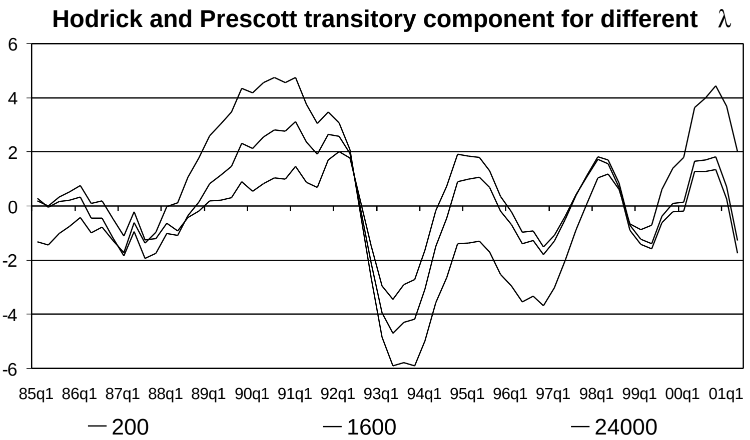

Figure 4.7 Estimated transitory component of the quarterly index of industrial production in the Euro-zone applying the Hodrick and Prescott filter with $\lambda = 200$, $\lambda = 1600$ and $\lambda = 24000$.

4.6.2. Infinite sample properties of the Hodrick and Prescott filter

A detailed analysis of the properties of the Hodrick and Prescott filter has been carried out by King and Rebelo (1993). Those authors considered the special case where the sample size is infinite, so the reference series is $\{y_t\}_{t=-\infty}^{\infty}$. This hypothesis, although not meaningful for empirical purposes, allows simplification of the algebra involved by the application of the Hodrick and Prescott filter and to derive useful properties of the filter itself, on both the frequency and time domain.

If the reference series is $\{y_t\}_{t=-\infty}^{\infty}$ the matrix equation (4.7) can be written, in the lag notation, as:

$$y_t = M(L)g_t$$

where:

$$M(L) = 1 + \lambda(1-L)^2(1-L^{-1})^2$$

So the trend component is obtained by applying the linear filter $H_g(L) = M^{-1}(L)$ to the observed series:

(4.9) $$g_t = H_g(L)y_t = \frac{1}{1+\lambda(1-L)^2(1-L^{-1})^2}y_t$$

and the filter to obtain the transitory component is $H_d(L)$ given by:

(4.10) $$d_t = H_d(L)y_t = \left[1 - H_g(L)\right]y_t = \frac{\lambda(1-L)^2(1-L^{-1})^2}{1+\lambda(1-L)^2(1-L^{-1})^2}y_t$$

One important consequence of (4.10) is that $H_d(L)$ can be factorised as:

$$H_d(L) = (1-L)^4 H'_d(L)$$

so the application of the Hodrick and Prescott filter can render stationary series containing up to four unit roots.

King and Rebelo (1993) show that the lag polynomial in (4.9) is symmetric. In fact denoting by θ_1 and θ_2 the roots of $M(L)$ that satisfy $|\theta_i| < 1$, $H_g(L)$ can be written as:

$$H_g(L) = \frac{\theta_1\theta_2}{\lambda}\left[\sum_{j=0}^{\infty}(A_1\theta_1^j + A_2\theta_2^j)L^j + \sum_{j=0}^{\infty}(A_1\theta_1^j + A_2\theta_2^j)L^{-j}\right]$$

where $A_1 = [(1-\theta_2/\theta_1)(1-\theta_1)^2(1-\theta_1\theta_2)]^{-1}$ and A_2 is the complex conjugate of A_1. So, with an infinite sample, the Hodrick and Prescott filter is symmetric and induces no phase shift.

The frequency response function of the filter $H_d(L)$ introduced in (4.10) is a real number and is equal to:

$$\tilde{H}_d(\omega) = \frac{4\lambda(1-\cos(\omega))^2}{1+4\lambda(1-\cos(\omega))^2}$$

Figure 4.8 shows the gain of the Hodrick and Prescott filter for the transitory component, in the case of $\lambda = 1600$, together with the ideal high pass filter for quarterly data, for which the cut-off frequencies respect $|\omega^*| = \pi/16$. As the figure shows, for $\lambda = 1600$ the Hodrick and Prescott filter offers a good approximation to the ideal filter.

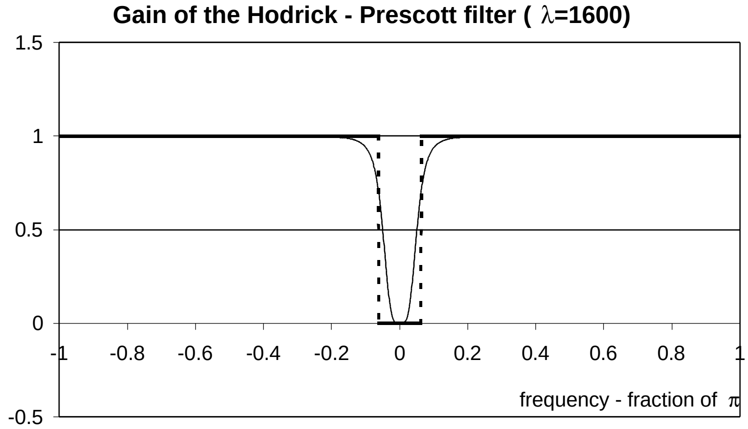

Figure 4.8 Gain of the Hodrick and Prescott filter for $\lambda = 1600$.

The same is not true for other values of λ: figure 4.9 plots the gain of the Hodrick and Prescott filter for the same values of λ used in figures 4.6 and 4.7. As the plot shows, for $\lambda = 200$ and $\lambda = 24000$ the approximation of the Hodrick and Prescott filter to the ideal filter for quarterly data is rather poor.

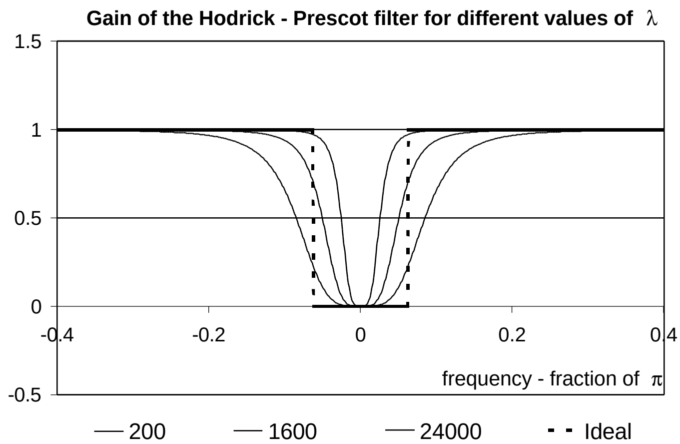

Figure 4.9 Gain of the Hodrick and Prescott filter for $\lambda = 200$, $\lambda = 1600$ and $\lambda = 24000$.

4.6.3. Optimality conditions for the Hodrick and Prescott filter

Harvey and Jaeger (1993) and King and Rebelo (1993) studied the conditions under which the Hodrick and Prescott filter performs as the optimal filter, in the sense that it minimises the mean squared error $MSE = (1/T)\Sigma_{t=1}^{T}(\hat{d}_t - d_t)^2$ where d_t is the "true" transitory component and $\hat{d}_t$ is its estimate. Suppose that the permanent and the transitory components follow two *ARMA* processes whose innovations are uncorrelated, that is:

(4.11) $A^g(L)g_t = M^g(L)\varepsilon_t^g$ and $A^d(L)d_t = M^d(L)\varepsilon_t^d$

where ε_t^g and ε_t^d are serially and mutually uncorrelated errors with variances, respectively, σ_g^2 and σ_d^2. The *AR* polynomial $A^g(L)$ for the permanent component can contain one (or more) unit roots.

Whittle (1963) shows that the optimal filter for d_t is given by:

(4.12) $$H_d^*(L) = \frac{\psi A^g(L)A^g(L^{-1})}{\psi A^g(L)A^g(L^{-1}) + (1-\psi)Q(L)}$$

where $\psi = \dfrac{\sigma_d^2}{\sigma_d^2 + \sigma_g^2}$ and $Q(L) = \dfrac{A^d(L)A^d(L^{-1})M^g(L)M^g(L^{-1})}{M^d(L)M^d(L^{-1})}$.

King and Rebelo (1993) study the conditions under which the Hodrick and Prescott filter (4.10) is coincident with the optimal filter (4.12). These conditions turn out to be very particular and sometimes in contrast with business cycle economic theory.

For example (4.10) is coincident with (4.12) if the *AR* and *MA* polynomials in (4.11) are given by:

$$A^g(L) = (1-L)^2 \qquad \text{and} \qquad M^g(L) = A^d(L) = M^d(L) = 1$$

that is the growth and transitory components are generated by:

$$\Delta^2 g_t = \varepsilon_t^g \qquad \text{and} \qquad d_t = \varepsilon_t^d$$

In this case the parameter λ is given by $\lambda = \sigma_d^2 / \sigma_g^2$. In their original paper Hodrick and Prescott (1980) argued that for quarterly data σ_d is approximately forty times bigger than σ_g, so that they suggest to choose $\lambda = (40)^2 = 1600$ for quarterly data.

The above mentioned optimality conditions are unlikely to be satisfied, as shown by Guay and St-Amant (1997). The most questionable assumptions are:

- Transitory and trend components are not correlated with each other. This implies that the growth and cyclical components of a time series are assumed to be generated by distinct economic forces, which is in contrast with some business cycle models (see Singleton (1988) for a discussion).

- The process is integrated of order two. This is often incompatible with priors on macroeconomic time series. For example, it is usually assumed that real GDP is integrated of order one or stationary around a breaking trend.

- The transitory component is white noise. This is also questionable. For example, it is unlikely that the stationary component of output is strictly white noise. King and Rebelo (1993) show that this condition can be replaced by the following assumption: an identical dynamic mechanism propagates changes in the trend component and innovations to the cyclical component. However, this condition is also very restrictive.

- The parameter controlling the smoothness of the trend component is appropriate. Economic theory provides little or no guidance to determine the ratio between σ_d^2 and

σ_g^2. While attempts have been made to estimate this parameter using maximum-likelihood methods (see Harvey and Jaeger (1993)) it appears difficult to estimate it with reasonable precision.

4.6.4. The Hodrick and Prescott filter in a finite sample

Section 4.6.2 presented the main properties of the Hodrick and Prescott filter on an infinite sample, that is when the series of interest is $\{y_t\}_{t=-\infty}^{\infty}$. Two major conclusions were drawn:

- the Hodrick and Prescott filter is symmetric, so it does not induce any phase shift at any frequency;
- the gain of the Hodrick and Prescott filter approximates the gain of the ideal high pass filter, provided the choice of the parameter λ is accurate (see figures 4.8 and 4.9).

These conclusions are no longer true if we consider the Hodrick and Prescott filter applied to a finite sample, that is (4.8). In this case, in fact, the filter weights change according to the period t in which we are estimating the trend. The filter is no longer symmetric and a phase shift is induced at every frequency, even if it is negligible except for the first and the last observations in the sample.

Moreover for these first and last observations the filter does not approximate anymore the ideal high pass filter. Figure 4.10 shows the gain for the Hodrick and Prescott filter calculated in $t=2$, $t=3$ and $t=4$.

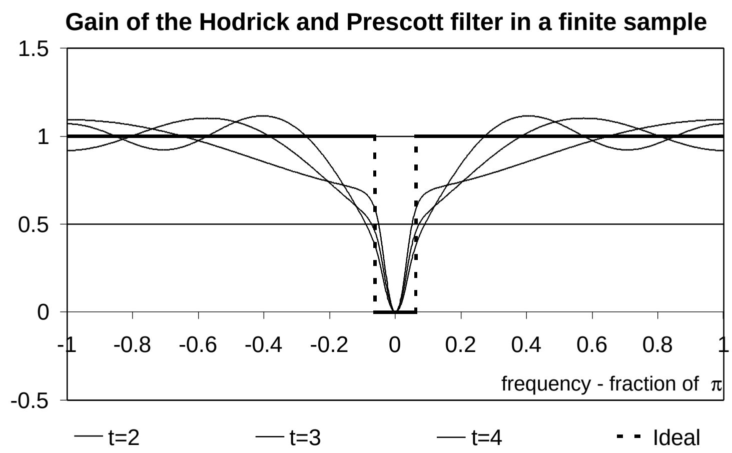

Figure 4.10 Gain of the Hodrick and Prescott filter with $\lambda=1600$ in a finite sample for $t=2$, $t=3$ and $t=4$.

4.7 The Baxter and King filter

The approach of Baxter and King (1995) to cycle extraction is essentially an attempt to approximate, through a finite terms centred moving average, the ideal band pass filter introduced in section 3.2. The approach is based on the definition of the business cycle proposed by Burns and Mitchell (1946), according to which the business cycle consists of fluctuations lasting no less than six and no more than thirty two quarters.

We will present Baxter and King methodology in two steps: first we will derive the approximation to the ideal low pass filter and then we will extend the method to obtain the approximation to the ideal high pass and band pass filter. In particular, the derivation of the high pass filter will allow for a comparison between the methodology of Baxter and King and the methodology of Hodrick and Prescott.

4.7.1. The low pass and high pass filter in the Baxter and King approach

As was pointed out in section 3.2, the ideal low pass filter, denoted by $H^*_{LP}(L)$, is characterised by a frequency response function whose gain $G^*_{LP}(\omega)$ and phase $\Psi^*_{LP}(\omega)$ fulfil the following conditions:

$$\begin{aligned} G^*_{LP}(\omega) &= 1 \quad \text{for every } \omega \text{ such that } |\omega| \le \omega^* \\ G^*_{LP}(\omega) &= 0 \quad \text{for every } \omega \text{ such that } |\omega| > \omega^* \\ \Psi^*_{LP}(\omega) &= 0 \quad \text{for every } \omega \end{aligned} \tag{4.13}$$

where ω^* is the cut-off frequency. For example if the ideal low pass filter is meant to extract the trend component from an economic quarterly series then, according to Burns and Mitchell (1946), the cut-off frequency corresponds to fluctuations of period $T = 32$ quarters and so we have $\omega^* = 2\pi/T = \pi/16$.

The time domain representation of the ideal filter $H^*_{LP}(L)$ is an infinite terms symmetric moving average:

$$H^*_{LP}(L) = \sum_{j=-\infty}^{\infty} b_j L^j \tag{4.14}$$

where $b_j = b_{-j}$. The coefficients b_j in (4.14) can be obtained by computing the inverse Fourier transform of the step function (4.13):

$$b_j = \int_{-\pi}^{\pi} G^*_{LP}(\omega) e^{i\omega j} d\omega = \int_{-\omega^*}^{\omega^*} e^{i\omega j} d\omega \tag{4.15}$$

Evaluating the integral (4.15), the coefficients b_j for the ideal low pass filter are given by:

$$\begin{aligned} b_0 &= \omega^*/\pi \\ b_j &= \sin(j\omega^*)/j\pi \quad \text{for } j = 1,2,... \end{aligned} \tag{4.16}$$

Of course the application of (4.14) requires an infinite sample period, so the main target of Baxter and King (1995) was to approximate (4.14) via a *finite* terms symmetric moving average like:

$$H^k_{LP}(L) = \sum_{j=-k}^{k} a_j L^j \tag{4.17}$$

where $a_j = a_{-j}$; so the moving average (4.17) involves only $2k+1$ terms. The weights a_j must be chosen so that they fulfil, for any *k*, the following constraints:

- the gain $G_{LP}^k(\omega)$ of $H_{LP}^k(L)$ is equal to one at frequency zero, that is $G_{LP}^k(0)=1$, like for the ideal low pass filter. This condition requires that the coefficients a_j in (4.17) respect $\Sigma_{j=-k}^k a_j = 1$;

- the gain $G_{LP}^k(\omega)$ is such that the following quadratic loss function:
$$Q = \int_{-\pi}^{\pi} \left| G_{LP}^*(\omega) - G_{LP}^k(\omega) \right|^2 d\omega$$
is minimised.

Baxter and King (1995) show that the a_j that fulfil these conditions are given by:

(4.18) $$a_j = b_j + \theta \quad \text{for } j = -k,\dots,k$$

where b_j is the *j*th coefficient of the infinite moving average (4.14) and θ is an additive constant so that $\Sigma_{j=-k}^k a_j = 1$, that is:

(4.19) $$\theta = \left(1 - \Sigma_{j=-k}^k b_j\right) \Big/ (2k+1)$$

So the best approximation to an ideal low pass filter is obtained by truncating (4.14) between $\pm k$ and by adding a constant so that the coefficients sum to one.

A crucial aspect for empirical applications is clearly the choice of *k* which determines the length of the truncated moving average. Baxter and King (1995) plotted, for quarterly data, the gain $G_{LP}^k(\omega)$ for different values of *k* against the gain $G_{LP}^*(\omega)$ of the ideal low pass filter; although it is clear that the larger *k* is, the better $G_{LP}^k(\omega)$ approximates $G_{LP}^*(\omega)$, the authors conclude that no sensible improvement is obtained beyond $k = 12$.

Figure 4.11 plots the gain of the Baxter and King low pass filter for $k=4$, $k=12$ and $k=20$ together with the gain of the ideal low pass filter. As the plot shows, the approximation is very poor for $k=4$ but it is rather acceptable for $k=12$ and even more for $k=20$.

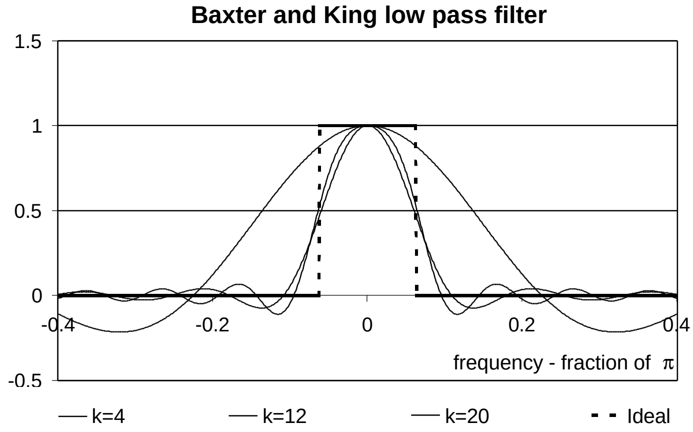

Figure 4.11 Gain of the Baxter and King low pass filter for $k=4$, $k=12$ and $k=20$ (for quarterly data and cut-off frequency $\omega^*=\pi/16$).

So in the Baxter and King approach the trend component g_t of the series y_t is obtained by:

$$g_t = H_{LP}^{12}(L)y_t = \sum_{j=-12}^{12} a_j y_{t-j}$$

where the coefficients a_j are given by (4.18) and the cut-off frequency is $\omega^*=\pi/16$. On the other end, an approximated high pass filter would extract from y_t the transitory component d_t, that is:

(4.20) $$d_t = H_{HP}^{k}(L)y_t = \sum_{j=-k}^{k} a'_j y_{t-j}$$

The derivation of this approximated high pass filter is straightforward. In fact we have:

(4.21) $$d_t = y_t - g_t = (1 - H_{LP}^{k}(L))y_t$$

and the weights a'_j in (4.21) are easily derived from the weights a_j in (4.18) by setting:

$$a'_0 = 1 - a_0$$
$$a'_j = -a_j \quad \text{for } j=-k,...,k, \quad j \neq 0$$

So the high pass Baxter and King filter (4.21) can be compared to the Hodrick and Prescott filter introduced in section 4.6. Figure 4.12 plots the gain of the Hodrick and Prescott filter compared to the high pass version of the Baxter and King filter for $k=12$. As the plot shows, there are no relevant differences between these two filters.

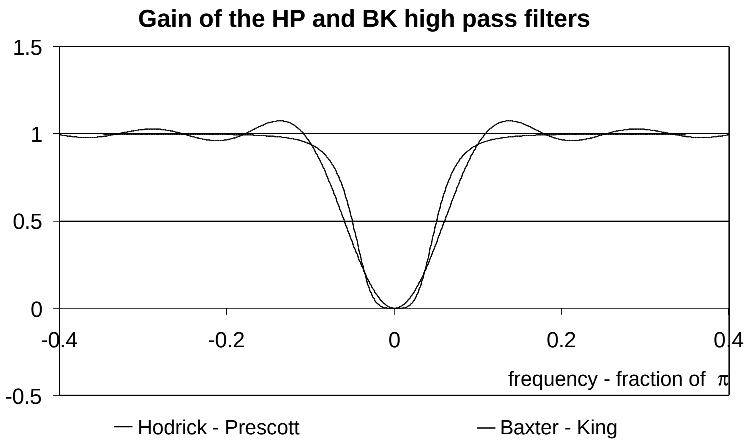

Figure 4.12 Gain of the Hodrick and Prescott filter (for quarterly data and $\lambda = 1600$) compared to the gain of the Baxter and King high pass filter (for quarterly data, $k = 12$ and cut-off frequency $\omega^* = \pi/16$).

4.7.2. The band pass Baxter and King filter

The most important contribution of Baxter and King (1995) paper is the derivation of a band pass filter to estimate directly the cyclical component c_t. Burns and Mitchell (1946) define the business cycle as fluctuations in some macroeconomic series lasting no less than 6 and no more than 32 quarters. Thus the natural goal of Baxter and King approach is to develop a band pass filter, that extracts only the components whose frequency lie in a particular range $[\omega_1^*, \omega_2^*]$, approximating the ideal band pass filter introduced in section 3.2. If the cut-off frequencies ω_1^* and ω_2^* are chosen appropriately then the cyclical component c_t is extracted by applying to y_t the following moving average:

(4.22) $$c_t = H_{BP}^k(L) y_t = \sum_{j=-k}^{k} a_j'' y_{t-j}$$

whose coefficients are given by:

(4.23) $$a_j'' = a_{2j} - a_{1j} + \theta_2 - \theta_1$$

where a_{1j} and a_{2j} are the *j*th weights, given by (4.18), of the low pass filters with cut-off frequencies ω_1^* and ω_2^* respectively, and θ_1 and θ_2 are the corresponding correction terms, defined by (4.19). It is easy to show that the coefficients a_j'' in (4.23) sum to zero.

As mentioned before, the choice of ω_1^* and ω_2^* depends from the range of periodicity we want to extract and from the frequency of the original data (annual, quarterly or monthly). For example to extract from quarterly data fluctuations in the range from 1.5 to 6 years we would fix $\omega_1^* = \pi/16$ and $\omega_2^* = \pi/3$.

Figure 4.13 shows, for quarterly data, the gain of the Baxter and King band pass filter for different values of *k* together with the ideal band pass filter. Like in the previous case, the approximation to the ideal filter is rather poor for small values of *k* and improves when *k* is increased. Once again no sensible improvement is obtained beyond $k = 12$, which is then chosen as a reference value for applications.

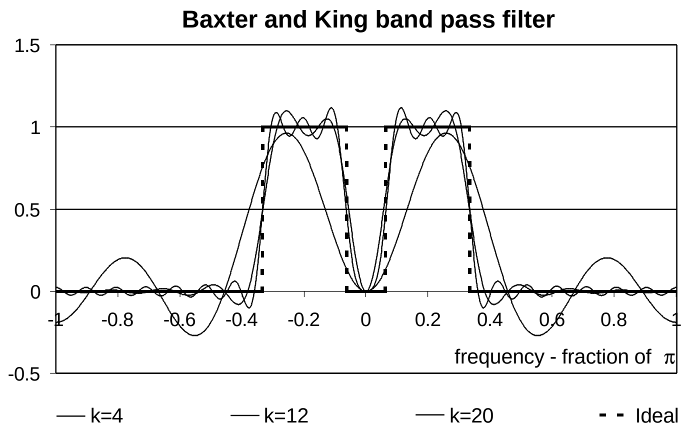

Figure 4.13 Gain of the Baxter and King band pass filter for $k = 4$, $k = 12$ and $k = 20$ (for quarterly data and cut-off frequencies $\omega_1^* = \pi/16$ and $\omega_2^* = \pi/3$).

As an example, figure 4.14 shows the weights of Baxter and King moving average (4.22) for $k = 12$, $\omega_1^* = \pi/16$ and $\omega_2^* = \pi/3$.

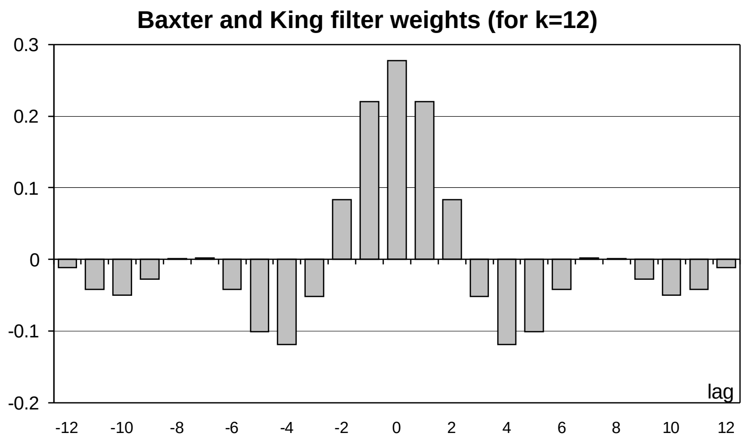

Figure 4.14 Weights of the Baxter and King band pass filter (for quarterly data, $k = 12$ and cut-off frequencies $\omega_1^* = \pi/16$ and $\omega_2^* = \pi/3$)

4.7.3. Some comments on the Baxter and King filter

When using the Baxter and King filter, *k* observations are lost at the beginning and the end of the sample period, according to the required degree of approximation to the ideal filter. As we showed in sections 4.7.1 and 4.7.2, the common choice is $k = 12$ for quarterly data. In order to reduce the loss of data at the beginning and at the end of the sample, truncated versions of the filter can be used.

Alternatively, it is possible to forecast and backcast the series before applying the filter so as to use the complete moving average. Anyway the reliability of forecasts over such a long period is quite low. Moreover if this forecast is made with univariate *ARIMA* models it is well known that they will not correctly capture turning points. So the use of truncated moving averages seems to be preferable, although they are subject to the same shortcomings that were pointed out for the finite sample version of the Hodrick and Prescott filter in section 4.6.4.

It is also important to observe that the Baxter and King filter is the only univariate method explicitly designed for the direct estimation of the cyclical fluctuation. Modifying the cut-off frequencies ω_1^* and ω_2^* it is also possible to isolate other specific cyclical components like, for example, Juglar and Kuznets cycles.

Another important feature of this filter is that, thanks to its band pass structure, it can deal with seasonally adjusted series as well as with unadjusted ones. In section 1 we saw that business cycle analysis is typically performed with seasonally adjusted data since they have a more regular behaviour which allows for an easier interpretation of the short term movements of the economy. Even if the estimation of the cyclical component could be performed in an equivalent way by applying the Baxter and King filter to both seasonally adjusted and unadjusted series, in practice, things are different. This is due to the fact that, in practice, economic series are often short and unadjusted data are often too erratic or noisy. Applying the same filter to seasonally adjusted and unadjusted (short) series does not produce the same results, see Astolfi et al. (2001).

4.8 The Beveridge and Nelson decomposition

Beveridge and Nelson (1981) were the first to propose a model based decomposition method of an integrated time series, which provides a convenient way to estimate its permanent and its transitory components. They show that any *ARIMA*(*p*,1,*q*) process can be represented as the sum of a stochastic trend plus a stationary component, where a stochastic trend is defined to be a random walk, possibly with drift. This representation is most easily obtained for an *ARIMA*(0,1,1) model. So suppose that Δy_t is a *MA*(1) process, so that $\Delta y_t = e_t + be_{t-1}$, where e_t is white noise and $|b| < 1$. Then:

(4.24) $$y_t = y_{t-1} + e_t + be_{t-1}$$

Solving equation (4.24) recursively and assuming $y_0 = e_0 = 0$ we obtain:

$$y_t = \sum_{j=1}^{t} e_j + b\sum_{j=1}^{t-1} e_j$$

and therefore:

(4.25) $$y_t = (1+b)\sum_{j=1}^{t} e_j - be_t$$

Equation (4.25) gives the decomposition of the series y_t into the trend and the transitory components, which are given by, respectively:

$$g_t = (1+b)\sum_{j=1}^{t} e_j \qquad \text{and} \qquad d_t = -be_t$$

An alternative expression for the trend component is:

$$g_t = g_{t-1} + (1+b)e_t .$$

Evidently g_t is a random walk without drift and d_t is a stationary process. Note that innovations in the two components are both proportional to e_t, i.e. they are perfectly correlated; if b>0, the correlation is –1 and if b<0 the correlation is +1.

The same result can be obtained for a general *ARIMA*(p,1,q) process. An *ARIMA*(p,1,q) process can be written as

$$a(L)\Delta y_t = f + b(L)e_t$$

where f is a constant and $a(L) = 1 - \Sigma_{j=1}^{p} a_j L^j$ and $b(L) = 1 + \Sigma_{j=1}^{q} b_j L^j$ are two lag polynomials of order p and q, respectively, whose roots lie outside the unit circle. Inverting $a(L)$ the model can be rewritten as an infinite *MA* process:

(4.26) $$\Delta y_t = h + c(L)e_t$$

where $h = f \big/ \Sigma_{j=0}^{p} a_j$ and $c(L) = b(L)/a(L)$. The right term in equation (4.26) can be decomposed in:

$$\Delta y_t = h + c(1)e_t + [c(L) - c(1)]e_t$$

that can be written as:

(4.27) $$\Delta y_t = h + c(1)e_t + (1-L)c^*(L)e_t$$

where $c^*(L)$ is the lag polynomial such that $(1-L)c^*(L) = c(L) - c(1)$. Multiplying both sides of (4.27) by $(1-L)^{-1}$ and assuming $y_0 = e_0 = 0$ we obtain:

(4.28) $$y_t = ht + c(1)\sum_{j=1}^{t} e_j + c^*(L)e_t$$

Equation (4.28) gives the Beveridge and Nelson decomposition for the series y_t. The permanent component is given by:

(4.29) $$g_t = ht + c(1)\sum_{j=1}^{t} e_j$$

and can be viewed as the sum of a "deterministic trend" *ht* and a "stochastic trend" $c(1)\Sigma_{j=1}^{t} e_j$. The transitory component is given by:

(4.30) $$d_t = c^*(L)e_t$$

As in the ARIMA(0,1,1) case, the innovations in the trend and in the cyclical components are both proportional to e_t and thus are perfectly correlated.

Some general remarks apply to the Beveridge and Nelson decomposition:

- Forecasting the series of interest is equivalent to forecasting the permanent component given by the Beveridge and Nelson decomposition since the transitory component has zero mean. So it is impossible to forecast the transitory component and, in particular, the turning points.

- The transitory component is defined as a cumulative sum of shocks so the cyclical pattern is generated by a diffusion mechanism like the one described by Frisch (1933). This implies that the cycle obtained using the Beveridge and Nelson decomposition can significantly differ with respect to the one obtained using mechanical approaches, like the Hodrick and Prescott filter.

- No a priori definition of the characteristics of the cyclical component is required by the Beveridge and Nelson approach. All the parameters involved in the decomposition are estimated from the available dataset so the decomposition does not depent upon “external” parameters linked, for example, to the smoothness of the trend or to the frequency range of the cyclical component.

- The estimated permanent and transitory components depend heavily on the particular *ARIMA* model which is fitted to the data. It is well known that the identification of an *ARIMA* model is a quite subjective issue. The Beveridge and Nelson is thus characterised by a certain degree of arbitrariness and different decompositions can be consistent with the same dataset.

4.9 The unobserved components decomposition of Harvey

The unobserved components approach, has been introduced by Harvey (1985). The main idea is that each economic time series y_t is the result of some unobserved components:

$$y_t = g_t + c_t + \varepsilon_t$$

which is the decomposition model proposed at the beginning of this section. With this approach we can also treat non-seasonally adjusted series by using the more general model

$$y_t = g_t + c_t + s_t + \varepsilon_t$$

where s_t is the seasonal component.

For sake of simplicity we focus on the model for seasonally adjusted data. The key hypothesis is that all the component are stochastic and generated by independent processes. So the Harvey decomposition is based upon the hypothesis that trend and cycle have a separate dynamic structure. Thus they are supposed independent, at least in the basic version of the model, in contrast with the Beveridge and Nelson decomposition presented in the previous section.

The main advantage of this model is that it can deal with structural breaks via an adequate formulation of the trend generating mechanism. It is also important to observe that, since c_t is assumed to be generated by a stochastic process of zero mean and finite variance, namely $c_t \sim N(0,\Omega)$ forecasting y_t is equivalent to forecasting the trend component g_t.

4.9.1. Alternative models for the trend component

The general model for the trend component g_t proposed by Harvey is:

$$g_t = g_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

where η_t and ζ_t are orthogonal white noise. So that, $\mathrm{Cov}(\eta_t, \zeta_t) = 0$ for all *t*, *s*.

It is important to note that η_t gives the possibility to the trend to fluctuate around its deterministic path. By contrast, ζ_t affects the slope of the trend.

It is possible to show that this representation of the trend g_t corresponds to an *ARIMA* model integrated of order two and respecting some particular constraints. As special cases of this representation, when $\sigma_\eta^2 = \sigma_\zeta^2 = 0$, the trend g_t is purely deterministic:

$$g_t = g_{t-1} + \beta = g_0 + \beta t$$

When only $\sigma_\zeta^2 = 0$, g_t is a random walk with drift of the form:

$$g_t = g_{t-1} + \beta + \eta_t$$

Finally if $\sigma_\eta^2 = 0$, the trend is an integrated process of order two where its second difference is equal to a white noise. This formulation of g_t is commonly referred as "slowly moving smooth trend".

The general formulation proposed by Harvey allows a large variety of trend specifications with different characteristics. This can be viewed as one of the main advantages of this approach, which is able to deal with very different growth typologies.

On the other hand, as in the Beveridge and Nelson decomposition, the cyclical component is supposed to be a zero mean stationary process. So, on the long run, forecasting the series is equal to forecasting the trend component.

4.9.2. The cyclical component

In the Harvey approach (1985-1989) the cycle c_t is modelled on the basis of a linear process with the possibility of displaying some more or less regular fluctuations. The general model can also deal with some asymmetries in the fluctuations.

The general form of the linear process is given by:

(4.31) $$\begin{bmatrix} c_t \\ c_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} c_{t-1} \\ c_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}$$

where κ_t and κ_t^* are mutually uncorrelated white noise errors with the same variance.

The parameter ρ, satisfying $0 \leq \rho \leq 1$, is the amplification factor insuring the stability of the cycle and λ, satisfying $0 \leq \lambda \leq \pi$, defines the frequency of the cycle corresponding to a period of $2\pi/\lambda$. c_t^* is a technical variable dual of the cycle c_t.

If κ_t and κ_t^* are set equal to 0, the cycle is then deterministic and is given by:

(4.32) $$c_t = \rho^t(\alpha \cos \lambda t + \beta \sin \lambda t)$$

where $\alpha = c_0$ and $\beta = c_0^*$. If $\rho = 1$ the cyclical movements are constants and if $\rho < 1$ the cyclical fluctuations tend to decrease.

Harvey proposes a stochastic version of the deterministic cycle. The deterministic cycle is a linear combination of trigonometric functions and it doesn't appear really adequate to describe economic fluctuations. The presence of innovations in the stochastic version of the cycle allows a higher flexibility, which better describes the behaviour of economic variables.

The role of innovations is to define at each time new initial conditions for the process whereas such conditions are fixed and equal to α and β in the deterministic version.

The formalisation proposed by Harvey seems really similar to the impulse propagation mechanism described by Frisch (1933). The cyclical components of a series is the result of a random chronology of impulses of a given variance and of a propagation mechanism characterised by a virtual duration and amplification factor. The resulting component is characterised by expansion and recession phases which can be asymmetric, thus displaying a more complex pattern with respect to the deterministic cycle in (4.32).

The model (4.31) can be viewed has an *AR*(1) model and can be written in the form:

$$\begin{bmatrix} c_t \\ c_t^* \end{bmatrix} = \begin{bmatrix} 1-\rho L\cos\lambda & \rho L \sin\lambda \\ -\rho L \sin\lambda & 1-\rho L\cos\lambda \end{bmatrix}^{-1} \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}$$

So the representation for the cycle c_t is given by:

$$(1-2\rho L\cos\lambda + \rho^2 L^2)c_t = (1-\rho L\cos\lambda)\kappa_t + (\rho L \sin\lambda){\kappa^*}_t$$

The right side of this equation is a sum of moving average processes. For this reason it can be written in the following form:

$$(1-2\rho L\cos\lambda + \rho^2 L^2)c_t = (1-\theta L){\varepsilon^c}_t$$

Where ${\varepsilon^c}_t$ is the cyclical innovation and where θ is a function of the basic parameters of the model. It can be easily shown that the cycle c_t is a stationary *ARMA*(2,1) process if $\rho < 1$.

The parameters of the model are constrained by its structural specification so that the eigenvalues of the *AR* polynomial are in the complex region, which is the usual condition for a cyclical behaviour of an autoregressive process. When λ takes values 0 or π the process degenerates in $AR(1)$, since $\sin\lambda = 0$ in (4.31) and the cyclical component is simply described by:

$$c_t = \rho c_{t-1} + \kappa_t$$

It can be useful to express the cycle c_t, in its general form, as an infinite moving average process, which gives

$$c_t = \sum_{j=0}^{\infty} \psi_j \varepsilon^c{}_{t-j}$$

where the parameters ψ_j, after the estimation, take the form:

$$\psi_0 = 1 \qquad \text{and} \qquad \psi_j = \rho^j f_{\lambda,\rho}(j) \ \text{ for } j > 1$$

where $f_{\lambda,\rho}$ is a periodical function of ρ and λ of periodicity $2\pi/\lambda$.

The sequence of parameters ψ_j describes a decreasing cycle, if $\rho < 1$, of period $2\pi/\lambda$ and with amplitude defined by the parameters ρ and λ. The infinite moving average representation proposed above can be useful in understanding the propagation mechanism of the stochastic cycle. Each innovation is the origin of its own cycle. For a given time point *t* the observed value of c_t appear to be the result of a accumulation of elementary cycles generated by each past innovation. All elementary cycles have the same frequency, amplitude and degree of deceleration but they differ in terms of date, sign and size of the innovation generating them.

Model trend + cycle

$$y_t = g_t + c_t + \varepsilon_t$$
$$g_t = g_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

Restrictions	
$\sigma^2{}_\eta \neq 0, \sigma^2{}_\varsigma \neq 0, \sigma^2{}_\varepsilon \neq 0$	Local Linear Trend + stochastic cycle
$\sigma^2{}_\eta = 0, \sigma^2{}_\varsigma \neq 0, \sigma^2{}_\varepsilon \neq 0$	Slowly Moving Smooth Trend + stochastic cycle
$\sigma^2{}_\eta \neq 0, \sigma^2{}_\varsigma = 0, \sigma^2{}_\varepsilon \neq 0$	Random Walk With Drift + stochastic cycle
$\sigma^2{}_\eta = 0, \sigma^2{}_\varsigma = 0, \sigma^2{}_\varepsilon \neq 0$	Deterministic Trend + stochastic cycle

Figure 4.15: Trend specification in the structural decomposition of Harvey

4.9.3. Synthesis of alternative decomposition models

The unobserved component approach allows users to define a wide variety of models for both the growth and the cyclical components. Since for economic proposes only stochastic cycles can be judged of interest and since degenerated $AR(1)$ processes are of no interest since they do not produce any cyclical movement, the main differences on the decomposition is given by the specification of the growth component.

Figure 4.15 presents the most common used trend specifications with the associated restrictions and under the hypothesis that the cycle is represented by a stochastic, stationary, and non-degenerate *ARMA* process.

4.9.4. Extension of the model

In order to deal with some special economic features the general decomposition model presented above can be generalised mainly following three different lines:

a) *Possibility of defining cycles of different periodicity*

There is a wide range of economic fluctuations, which can characterise observed time series. Examples can be seasonal fluctuations and long term movements such as Kuznets cycles. So the cyclical part of the model can be appropriately adjusted to deal with such fluctuations. In this case, the decomposition scheme will be more complex involving, for example, a seasonal component s_t.

b) *Inclusion of intervention variables to treat structural breaks*

Structural changes can produce breaks in the trend evolution of the economy. To have a good representation of these structural changes, like for example the German unification, it is possible to include in the decomposition model four different types of intervention variables:

- *"Impulse intervention variables"* corresponding to an outlier in the irregular component ε_t;
- *"Step intervention variables"* describing a level shift in the series;
- *"Slope intervention variables"* representing a permanent change of the slope of the trend;
- *"Seasonal intervention variables"* representing a change in the seasonal pattern.

c) *Use of external variables to explain some features for which the simple univariate model cannot be adequate*

The inclusion of external variables can be justified when there are modification in the dynamic behaviour of the series which cannot be explained by its internal structure. This can be the case of external shocks such as a non-expect depreciation of the national currency (Italy 1994) or the oil crises (1974-1975).

By taking into account all the previous extensions, the model presented above can be rewritten in the following more general form:

$$y_t = g_t + \sum_{j=1}^{m} c_{jt} + \sum_{k=1}^{p} \sum_{\tau=0}^{q} \delta_{k\tau} x_{k,t-\tau} + \sum_{h=1}^{r} \theta_h w_{h,t} + \varepsilon_t$$

where there are *m* distinct cycles, p explanatory variables $x_{k,t}$, and *r* intervention variables $w_{h,t}$.

4.9.5. Estimation and testing

In order to estimate the unobserved component g_t and c_t the model is expressed in the usual state space model. The state equation is given by $z_t = Fz_{t-1} + \chi_t$ where z_t is the state vector including all the unobserved component, *F* is a so-called transition matrix and χ_t is the noise component. In our case the model takes explicitly the form:

$$z_t = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & -\rho\sin\lambda & \rho\cos\lambda \end{bmatrix} z_{t-1} + \begin{bmatrix} \eta_t \\ \varsigma_t \\ \kappa_t \\ \kappa_t^* \end{bmatrix}$$

and the variance covariance matrix of χ_t is diagonal. The measurement equation is given by $y_t = Hz_t + \varepsilon_t$ which can be written in the form:

$$y_t = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix} z_t + \varepsilon_t$$

The estimation of the unobserved components is obtained in a maximum likelihood framework via the Kalman Filter. The Kalman Filter gives a recursive estimate of the state vector z_t at each time *t* conditional on the available information.

Then a smoothing procedure allows to compute the expectation of the state vector z_t conditional on the available information over the overall observation period 1, ... ,*T*. This operation give the possibility of extracting at each time all unobserved component as well as their innovations.

Different diagnostics have been proposed to evaluate the adequacy of the model. The most commonly used one is the one step forecasting error defined as the variance of the difference between the observed values of y_t and the one step ahead forecast of the model. The square root of this measure gives the standard error associate to the model. The standardised errors defined as a the ratio between the one step forecast errors and the standard error of the model are then used to compute additional quality measures. Such measures are:

a) Normality statistics giving the possibility of identify the presents of unexplained outliers. Tests have been proposed by Bowman and Shenton (N_{BS}) and by Doornik and Hansen (N_{DH}) which are based on the first and fourth moment of the distribution. Those statistics are distributed as $\chi^2(2)$.
b) Absence of autocorrelation of the residuals based on the Box and Ljung test Q(p,q). Under the null hypothesis this statistics is distributed as $\chi^2(q)$ where *q* the number of parameters governing the stochastic component.

The statistics R^2_D is used to assess the forecasting performance of the one step forecasting performance of the estimated model in comparison with the benchmark model represented by a random walk drift given by $y_t = y_{t-1} + \beta + \eta_t$. If R^2_D takes positive values this means that the estimated model performs better than the benchmark one.

4.9.6. Some remarks

The unobserved components decomposition of Harvey presents some operational difficulties, both with regard to the parameters estimation and the extraction of the cyclical and trend components. These require, in fact, a state space representation of the model, and the extraction of the components is made using the Kalman filter.

The fact that the components are supposed to be uncorrelated can be viewed as an obstacle in particular in the short run when it can be likely to assume a correlation between growth and cyclical movements as proposed in the Beveridge and Nelson decomposition.

The main advantage of this approach is represented by its flexibility and by the possibility of using dummies and external variables to deal with events not easily treated in a purely univariate context. Moreover the fact that this approach can treat seasonal series can be viewed as an important improvement with respect to other approaches. In effect, seasonality, cyclical movements and trends can be simultaneously identified and extracted within the same procedure which ensures an internal consistency which does not characterise alternative decomposition approaches.

5. An application of univariate cycle extraction techniques to some Euro-zone data

In this section we propose a simple comparative analysis of different cycles obtained by using alternative extraction techniques.

5.1 Description of the data

The reference series is the (seasonally adjusted) industrial production index for the Euro zone. The seasonal adjustment has been obtained through the so-called "direct" approach, that is by aggregating unadjusted data from Member States and then removing the seasonal component from the total using TRAMO SEATS. The sample period is from January 1985 to March 2001. Figure 5.1 shows the evolution of the monthly industrial production index over the sample period.

The choice of industrial production index for this comparison can be easily explained. Even if the share of the industrial sector of total GDP is no more than 40%, the industrial fluctuations can explain more than 60% of the total GDP volatility.

This implies that the industrial production index is the main indicator when we want to analyse short-term cyclical movements. Moreover, this series is available on a monthly basis; this increases considerably its ability to describe cyclical fluctuations, for example with respect to the total GDP which is only available on quarterly basis.

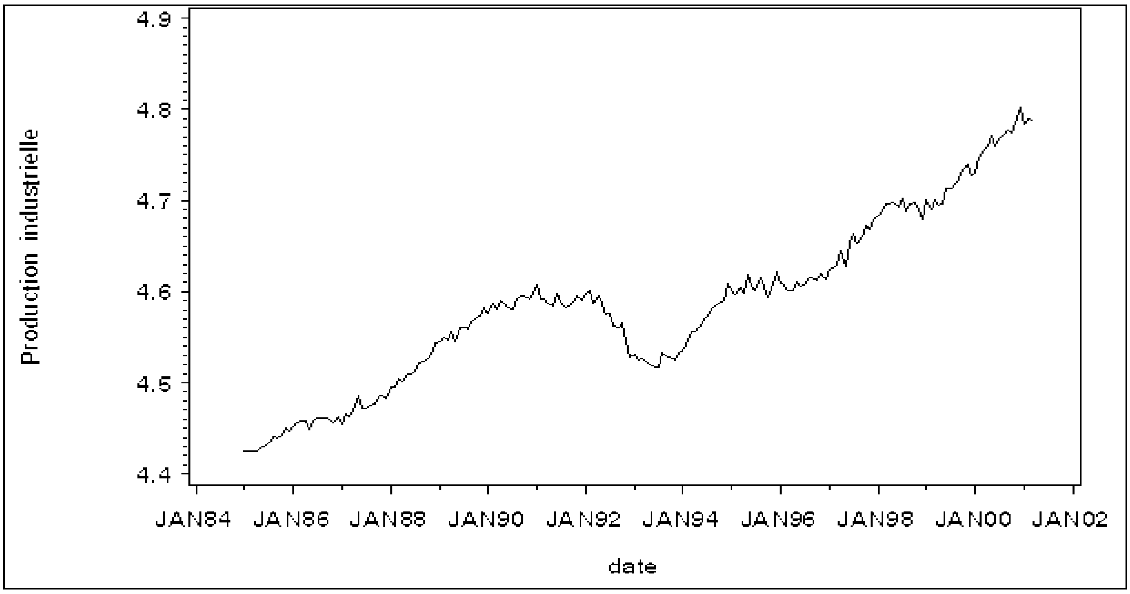

Figure 5.1 Industrial production index for the Euro zone

5.2 Euro zone cyclical fluctuation

It can be of high interest to derive the alternative estimations of the cyclical fluctuations in the industrial production index following some of the approaches presented in section 4 and to compare them in order to discover main similarities and differences. We focus our attention upon the following four approaches:

- Hodrick and Prescott: this approach is the most widely known and used in literature;
- Baxter and King: it is the only approach that directly estimates the cyclical component;
- Phase Average Trend: based on an iterative procedure, focuses on an accurate dating of turning points;
- Beveridge and Nelson decomposition: an approach that just considers the "stochastic" aspect of the problem.

Figures 5.2 to 5.5 show the alternative estimation of the Euro zone cycle obtained by using the above mentioned methods.

The comparison of figures 5.2, 5.3 and 5.4, shows quite clearly that the estimated cyclical component does not differ too much from one method to the other. Moreover, as was expected, the Baxter and King filter produces smoother estimates than the other two approaches. This is essentially due to the fact that the Baxter and King filter estimates directly the cyclical component c_t while the Hodrick and Prescott filter and the Phase Average Trend method produce estimates of the transitory component d_t which includes also the irregular part ε_t.

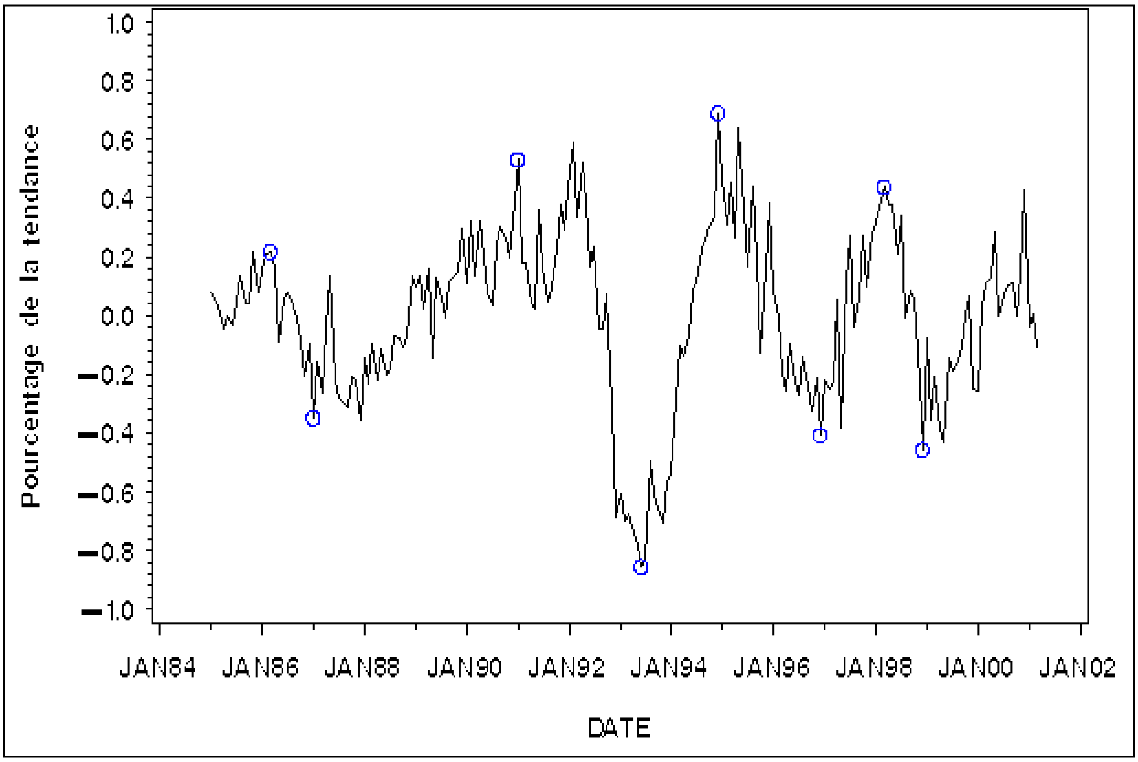

Figure 5.2 Hodrick and Prescott cycle of the Euro zone industrial production index

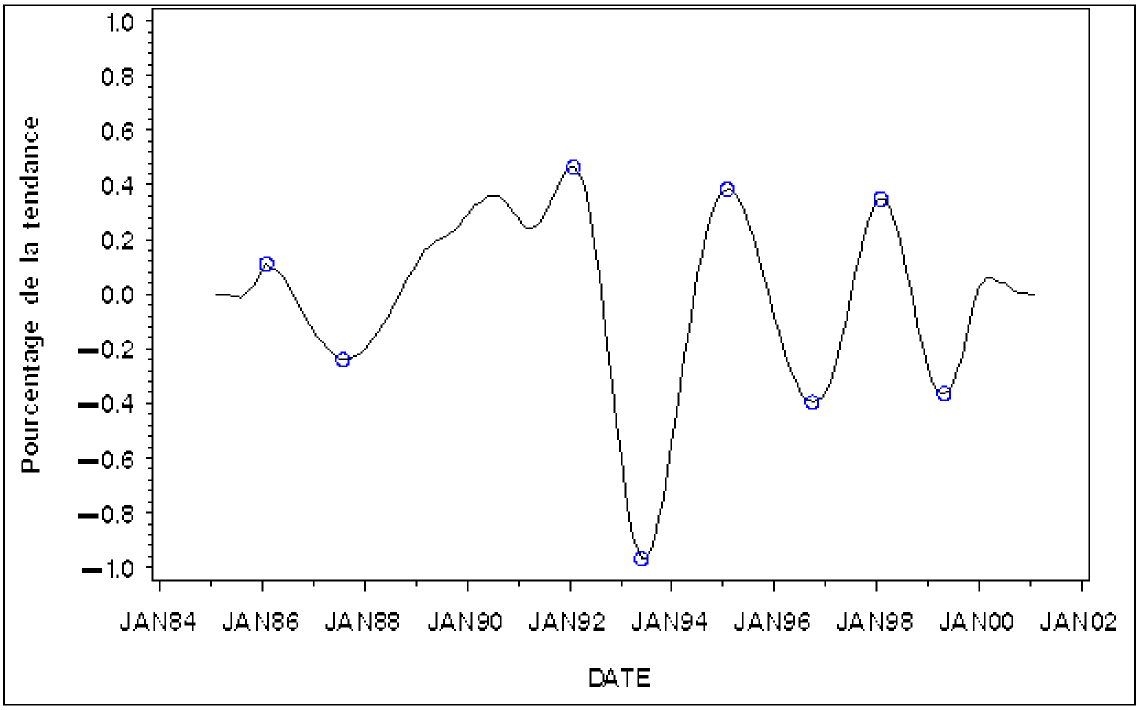

Figure 5.3 Baxter and King cycle of the Euro zone industrial production index

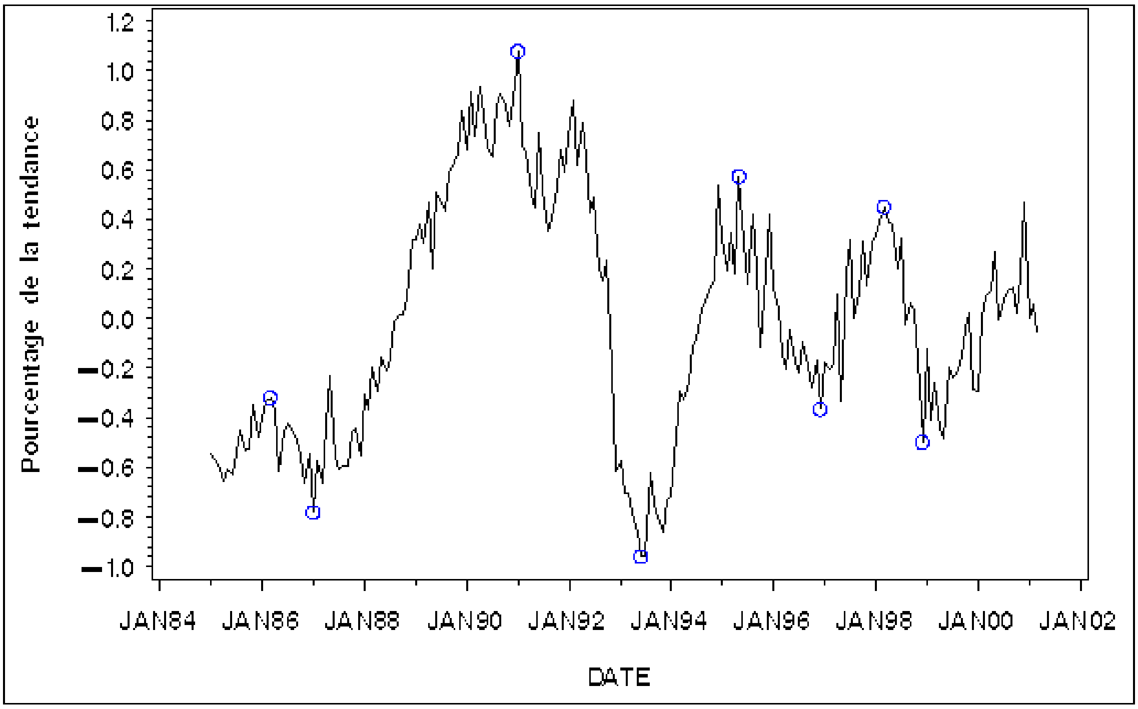

Figure 5.4 PAT (Phase Average Trend) cycle and turning point chronology of the Euro zone industrial production index

Regarding turning points, they are almost coincident in the series resulting from the HP and PAT methods. The turning points are slightly different in the series resulting from the BK approach.

These differences, which are not at all systematic, could be explained by the presence of the irregular component ε_t in the cycles derived from HP and PAT methods: this component may produce small distortions in the turning point detection which is not observed in the cycle derived from the BK filter.

Concerning the Beveridge and Nelson cycle, figure 5.5 shows five alternative estimates obtained by imposing different *ARIMA* structures to the industrial production series.

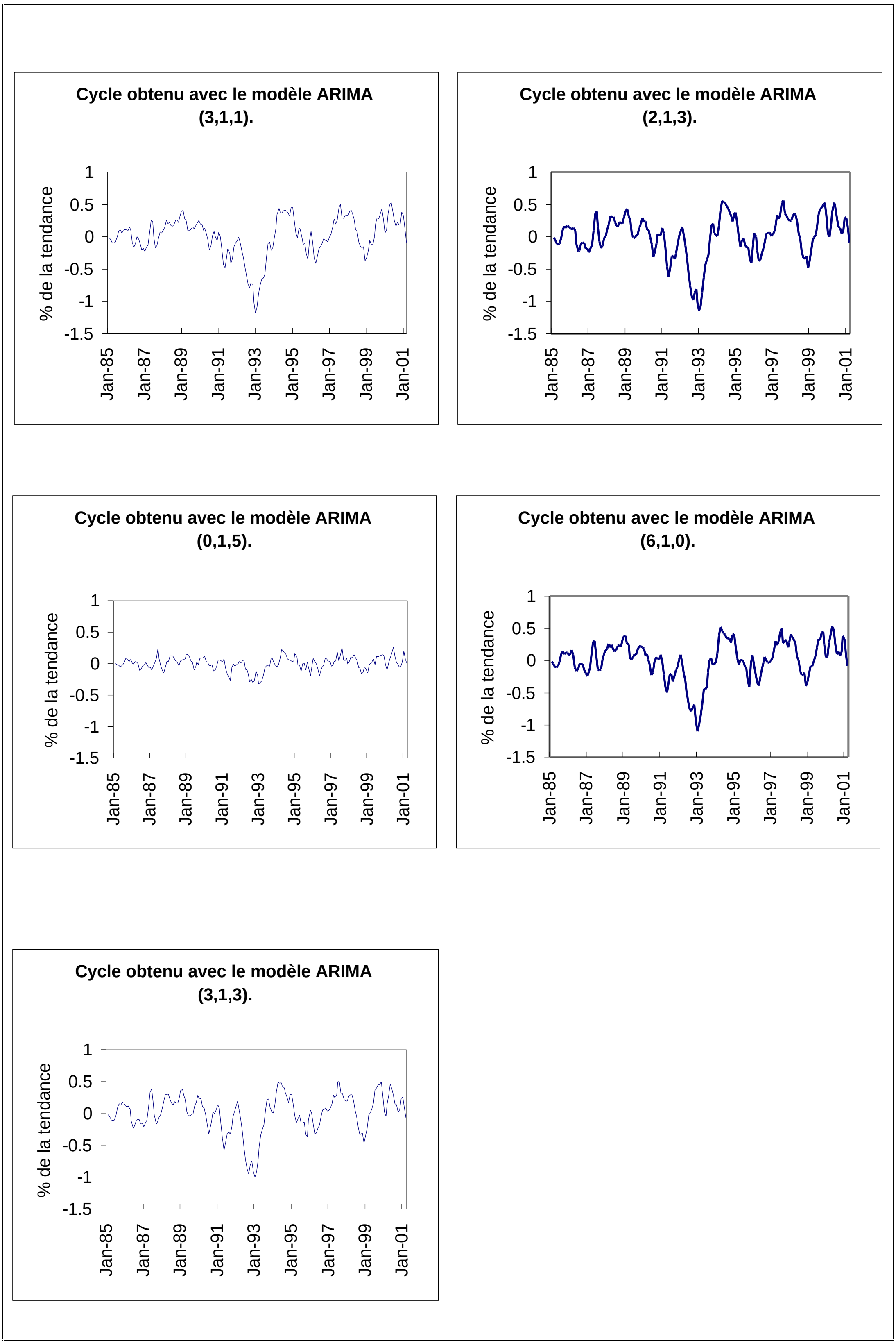

Figure 5.5 Alternative Beveridge and Nelson cycles of the Euro zone industrial production index

Clearly, Beveridge and Nelson cycles are quite different from the previous ones because they are showing a transitory component which is essentially different from the one that underlines purely mechanical approaches. The cycle is not based on some a priori specification, such as those proposed by the NBER, but it is determined by a cumulative sequence of stochastic shocks represented by an infinite moving average process. This implies that the Beveridge and Nelson cycles are generally less regular and less smoothed than the previous ones. This implies that quasi-periodical movements are much more evident in the cycles derived from mechanical filters than in cycles derived from the Beveridge and Nelson decomposition.

Looking at the figure 5.5, we can observe that the estimated cycles are quite similar, even if the *ARIMA* models on which they are based are different. The only exception is represented by the cycle estimated starting from an *ARIMA* model of order (0,1,5). This can be explained by the fact that this model is the only one in which the autoregressive part is missing. The cyclical component obtained by this model is significantly different from the others.

5.3 Simple comparative analysis

In this section, we propose a very simple graphical comparison of the different cycles outlined in section 5.2. For the Beveridge and Nelson decomposition only the cycle obtained from an *ARIMA*(3,1,1) has been considered.

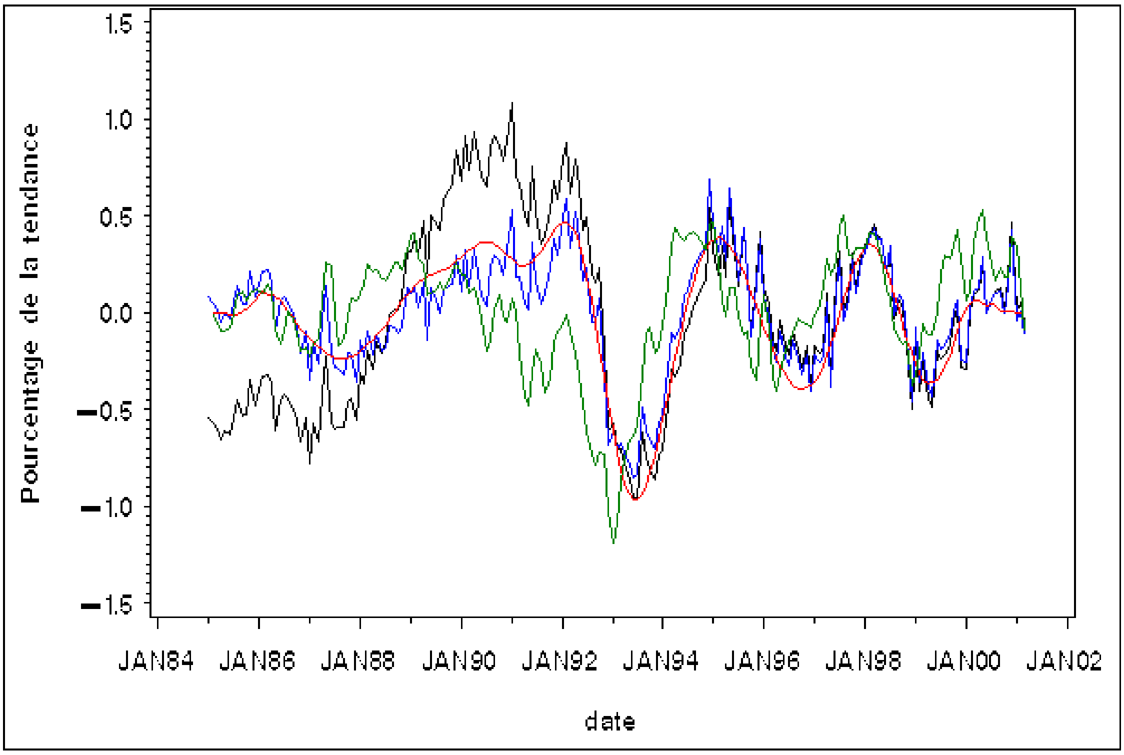

Figure 5.6 Comparison of cycle of Euro zone industrial production index obtained using the four methods

We can easily note as the Baxter and King cycle (red curve) and the Hodrick and Prescott cycle (blue curve) appear to be really similar. We could imagine the Baxter and King cycle as a smoothed version of the Hodrick and Prescott one.

The cycle derived with the Phase Average Trend method (black curve) is also quite similar to the two others already mentioned but only in the second half of the sample period. This could be produced, at least partially, by the extensive use of extrapolation techniques at the beginning and at the end of the series.

Finally, the Beveridge and Nelson cycle (green curve) is characterised by a similar pattern that the previous three, even if this evolution is less smoothed and characterised by a higher volatility which is completely consistent with its stochastic nature. Moreover, the Beveridge and Nelson turning point seem to be sometimes out of phase with respect to the Beveridge and Nelson cycle. For example, the Baxter and King cycle identifies a significative negative turning point in June 1993, which is also displayed by the Hodrick and Prescott and Phase Average Trend cycles.

Looking at the Beveridge and Nelson cycle, this negative turning point appears in January 1993 i.e. 6 months earlier. This phenomenon needs to be better investigated due to the fact that the Beveridge and Nelson decomposition is strongly dependent on the chosen ARIMA specification.

6. Multivariate methods for the estimation of potential output and output gap

In section 1 it has been shown how potential output is typically a constraint equilibrium position of the economic system which corresponds, in many cases, to a so called *steady state*. Steady state position is commonly defined as an equilibrium growth path without any tension. The constraints are defined by the absence of such tensions. Structural relationships such as Okun’s law, Phillips curve and NAIRU are just a few examples of the constraints to be imposed on the evolution of potential output.

Besides any economic consideration, the constraints can vary according to the time horizon of the analysis. In a long-run context the optimal utilisation of productive factors and the technological process play the most important role. By contrast in a shorter period the absence of inflationary pressures can be viewed as the main conditioning factor.

All univariate methods discussed in section 5 cannot deal by definition with all the above mentioned constraints. The potential output derived in that context is a sort of statistical trend obtained by smoothing the observed data.

The multivariate methods that we present in this section are an extension of the univariate ones; in many cases they allow the possibility of including structural economic constraints in the estimation process of potential output.

Such methods have been developed quite recently (starting from the end of 1980s, beginning of 1990s). This means that they are still in a testing phase where more research activities need to be done. This situation justifies the fact that only a limited number of empirical applications of such approaches are available at present (see for a good survey, Chagny and Döpke, 2001). Even if, the multivariate context appears to be most adequate for a robust estimation of potential output and the output gap, available results do not show clear evidence in favour of such methods in comparison with the univariate ones.

Moreover, economic theory is based on quite simple models which do not reflect the complexity of the economic system. The presence of non-linearity, unexpected structural and political changes as well as purely exogenous factors can affect the results obtained by using multivariate methods.

Finally, it must be also pointed out that, in the multivariate methods, the role of the subjective appreciation of analysts is increased since they have responsibility and the choice of the conditioning economic relationships to be included in the model.

6.1 The multivariate Hodrick and Prescott Filter

This method has been proposed by Laxton and Tetlow (1992) at the Bank of Canada. It is a generalisation of the univariate Hodrick and Prescott Filter already presented above. The main improvement of this filter is that it includes some structural relationships coming from the economic theory. Examples of economic relationships to be included are the Phillips curve, the Okun's Law and the NAIRU hypothesis. Moreover it is also possible to include information about the supply side equilibrium and the stage of the business cycle. In this sense this filter intends to solve one of the main shortcomings of the Hodrick and Prescott filter (see Chagny and Döpke 2001), i.e. the fact that it is defined without any reference to the economic theory.

6.1.1. Assessment of the filter

By assuming the presence of one conditioning economic relationship, the solution for the multivariate version of the HP filter is given by:

$$\min_{\{g_t\}_{t=1}^{T}} \sum_{t=1}^{T} \left[(y_t - g_t)^2 + \lambda_1 [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 + \lambda_2 \xi^2{}_t \right] \qquad (6.1)$$

for given λ_1 and λ_2 and where ξ_t is the residual from an estimated economic relationship:

$$y'_t = \beta g_t + \delta x_t + \xi_t$$

where y'_t is another variable depending on the growth component g_t and on an exogenous variable x_t and $\xi_t \sim N(0, \sigma^2 V)$.

This method has been used by the Bank of Canada (see Butler, 1996), by the Bank of New Zealand (see Conway and Hunt, 1997) and by OECD (1999). As in the univariate filter λ_1 and λ_2 reflect the weights of different components in the minimisation problem. They can consequently affect the degree of smoothness of the residual component d_t.

This can be shown by rewriting the first equation (6.1) in the following form:

$$\min_{\{g_t\}_{t=1}^{T}} \sum_{t=1}^{T} \left[\frac{1}{\sigma_0^2} (y_t - g_t)^2 + \frac{1}{\sigma_1^2} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 + \frac{1}{\sigma_2^2} \xi^2{}_t \right]$$

with $\lambda_1 = \dfrac{\sigma^2}{\sigma_1^2}$ and $\lambda_2 = \dfrac{\sigma_0^2}{\sigma_2^2}$.

In this case, the estimation of the growth component g_t is not simply a moving average on the observed value y_t but it includes also the elements coming from the economic relationship. The suitable λ_2 value is inversely proportional to the variance σ_2^2 of the residual component of the economic relationship. A higher value of λ_2 means that information added by the economic relationship is relevant.

An example of an application of the multivariate Hodrick and Prescott filter is given by De Brouwer (1998). Consider the estimation of potential output g_t by using the additional information coming from the Phillips curve, Okun's Law and capacity utilisation equations. We can then write the following minimisation problem:

$$L = \sum_{t=1}^{T}(y_t - g_t)^2 + \lambda_1 \sum_{t=2}^{T-1}(\Delta g_{t+1} - \Delta g_t)^2 + \sum_{t=1}^{T}\lambda_{2,t}\varepsilon_{\pi,t}^2 + \sum_{t=1}^{T}\lambda_{3,t}\varepsilon_{u,t}^2 + \sum_{t=1}^{T}\lambda_{4,t}\varepsilon_{cu,t}^2 \tag{6.2}$$

where $\varepsilon_{\pi,t}$, $\varepsilon_{u,t}$ and $\varepsilon_{cu,t}$ are the residuals in the regressions listed below and $\lambda_{2,t}$, $\lambda_{3,t}$ and $\lambda_{4,t}$ are possibly time-varying weights.

The residuals are obtained from the following three equations, which define, respectively, the Phillips curve, Okun's law and capacity utilisation.

- $\pi_t = \pi_t^e + A(L)(y_t - g_t) + \varepsilon_{\pi,t}$: which states that inflation is above expected value when output is above the non-accelerating inflation level of potential output.

- $u_t = nairu_t + B(L)(y_t - g_t) + \varepsilon_{u,t}$: which draws on Okun's law with the unemployment rate below the non-accelerating inflation rate of unemployment (NAIRU) when output is above its potential level.

- $cu_t = cu_t^g + C(L)(y_t - g_t) + \varepsilon_{cu,t}$: which draws on a partial indicator of supply capacity, stating that capacity utilisation is above trend when output is above potential.

Estimation of g_t minimising (6.2) is obtained through an iterative procedure.

Initial values for these equations are obtained by computing a preliminary estimation of g_t by using an univariate HP filter with λ=1600 as suggested by Hodrick and Prescott (1981-1997).

To estimate the three structural relations, additional hypothesis are needed in particular concerning the generating mechanism of expectations in the Phillips curve, the estimation of NAIRU and capacity utilisation to be made (see De Brouwer 1998).

An iterative procedure is used to obtain a final estimation of the potential output. As a first step the structural equations are estimated by using the results from the univariate Hodrick and Prescott filter as initial conditions. Then, the results from the structural equations are used in the multivariate Hodrick and Prescott model to compute again the potential output g_t. The output gap is consequently obtained and the structural equations are then re-estimated based on such new results. The loss function for the multivariate filter is then recomputed using the new results for the structural equations. This procedure will continue until the coefficients on the output gap satisfy specified convergence criteria, as described in Conway and Hunt (1997).

The problems of weighting different structural components in the loss function need also to be briefly discussed. The weights can be considered fixed or they can vary at each iteration. The main principle is that the weight should be inversely proportional to the size of the residuals from the corresponding equation. For example, using the output gap variance as reference criteria we can calculate the weight of the *j*th component as:

$$\lambda_j = \frac{\Sigma_{t=1}^{T}(y_t - g_t)^2}{\Sigma_{t=1}^{T}\varepsilon_{j,t}^2}$$

which can be recomputed at each iteration.

The estimation of the gap depends on the definition of the conditioning relationships and on the weighting scheme in the loss function. The way inflation expectations are generated, the choice of the narrow estimates as well as their relation with the expected inflation rate can considerably affect the final estimates of the output gap. Finally, it is also possible to envisage time varying weights λ_j which reflect the change of the relative importance of the conditioning relationships in the estimation period.

6.1.2. Some remarks

The multivariate HP filter is an improvement and an extension of the usual HP filter. Consequently, the same observations made to the HP filter can be applied to the multivariate one. Additionally, it is useful to underline how the results derived from the multivariate HP filter are really and strongly dependent on the subjective appreciation of the economic theory.

Neo-classical or Keynesian hypothesis on the structural relationships can produce really different patterns for the output gap. Moreover, the results of the filter are also conditioned by the choice of the convergence criteria and by the fact that nothing in principle can ensure that the iterative process described above will converge.

6.2 Multivariate Beveridge and Nelson decomposition

The Beveridge and Nelson decomposition in a multivariate framework was proposed by Evans and Reichlin (1994). The basic idea follows the definition, made by Beveridge and Nelson, of the trend component in terms of long-run forecast of the series of interest. By increasing the information set, i.e. by including additional explanatory variables, better forecasts can be made for the series of interest and thus, argue Evans and Reichlin, better estimates can be obtained of the permanent component of the series.

In a multivariate setting the variable y_t in (4.26) is replaced by two vectors: the first one, denoted by y_t contains k_1 $I(1)$ non-stationary variables and the second one, denoted by z_t, contains k_2 stationary variables. The first row of y_t contains the variable of interest (or its logarithm). Stacking in a single column vector the first differences of y_t and the levels of z_t, we can write the following Wold decomposition:

(6.3) $$\begin{bmatrix} \Delta y_t \\ z_t \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \begin{bmatrix} C_1(L) \\ C_2(L) \end{bmatrix} e_t = h + C(L)\, e_t$$

where $C_1(L)=\Sigma_{j=0}^{\infty}C_{1j}L^j$ and $C_2(L)=\Sigma_{j=0}^{\infty}C_{2j}L^j$ are two matrix lag polynomials and the errors e_t are white noise with $Var(e_t)=\Omega$.

The term $C(L)\,e_t$ in (6.3) can be decomposed into:

$$\begin{bmatrix}\Delta y_t\\ z_t\end{bmatrix}=\begin{bmatrix}h_1\\ h_2\end{bmatrix}+\begin{bmatrix}C_1(1)\\ 0\end{bmatrix}e_t+\begin{bmatrix}C_1(L)-C_1(1)\\ C_2(L)\end{bmatrix}e_t$$

which can be written as:

$$\begin{bmatrix}\Delta y_t\\ z_t\end{bmatrix}=\begin{bmatrix}h_1\\ h_2\end{bmatrix}+\begin{bmatrix}C_1(1)\\ 0\end{bmatrix}e_t+\begin{bmatrix}(1-L)C_1^*(1)\\ C_2(L)\end{bmatrix}e_t \tag{6.4}$$

where $C_1^*(L)$ is the solution to $(1-L)C_1^*(L)=C_1(L)-C_1(1)$.

Multiplying the first k_1 rows of (6.4) by $(1-L)^{-1}$ we obtain:

$$\begin{bmatrix}y_t\\ z_t\end{bmatrix}=\begin{bmatrix}h_1 t\\ h_2\end{bmatrix}+\begin{bmatrix}C_1(1)\\ 0\end{bmatrix}\sum_{j=1}^{t}e_j+\begin{bmatrix}C_1^*(1)\\ C_2(L)\end{bmatrix}e_t \tag{6.5}$$

which gives the multivariate Beveridge and Nelson decomposition for the vector $[y_t\ \ z_t]'$.

The rank of the matrix $C_1(1)$, which is less than or equal to k_1, is central in determining the long-run evolution of the variables y_t. If $C_1(1)$ is full rank then the k_1 variables in y_t are moved by k_1 independent random walks. If the rank of $C_1(1)$ is k_1-r this implies that there are *r* linear combinations of y_t which are stationary, that is the variables y_t are cointegrated and *r* is their cointegration rank.

Thus the rank of $C_1(1)$ gives the number of "common trends" that, following (6.5), determine the long-run evolution of the non-stationary variables y_t.

It must be noticed that an identification problem exists in (6.3) and it is impossible to associate one element (or a linear combination of elements) of e_t with a particular economic process. In fact suppose *S* is a non-singular matrix and a new set of innovations $\eta_t=S^{-1}e_t$ is introduced. Then (6.3) could be written as:

$$\begin{bmatrix}\Delta y_t\\ z_t\end{bmatrix}=h+C(L)\,SS^{-1}e_t=h+C(L)\,S\eta_t$$

So the identification of the parameters in (6.3) require a normalisation constraint. The usual strategy is choosing *S* so that:

$$\begin{bmatrix}C_{10}\\ C_{20}\end{bmatrix}=I$$

Quah and Vahey (1995) show an example where the normalisation is based on priors coming from the economic theory.

6.3 The common trends - common cycles decomposition of Vahid and Engle

Let y_t denote a vector of n $I(1)$ variables, so that Δy_t is stationary. We can write the Wold representation of Δy_t:

(6.6) $$\Delta y_t = h + C(L)\, e_t$$

where e_t are white noise with covariance matrix Ω. Equation (6.6) is analogous to (6.3) in the case that no additional stationary variables z_t are considered. Furthermore we assume $h = 0$ for ease of notation: this implies no deterministic trends in the levels of y_t. In (6.6) the lag matrix polynomial $C(L) = \Sigma_{j=0}^{\infty} C_j L^j$ satisfies the normalisation constraint $C_0 = I_n$ discussed in section 6.2.

The right term in (6.6) can be expanded as follows:

(6.7) $$\Delta y_t = C(1)\, e_t + (1-L)C^*(L)e_t$$

By integrating (6.7) we obtain the multivariate Beveridge and Nelson decomposition for y_t, introduced in section 6.2, that is:

(6.8) $$y_t = C(1)\sum_{j=1}^{t} e_j + C^*(L)e_t$$

where the permanent components of y_t are given by:

(6.9) $$g_t = C(1)\sum_{j=1}^{t} e_j$$

and the transitory components by:

(6.10) $$d_t = C^*(L)e_t = \sum_{j=0}^{\infty} C_j^* L^j e_t$$

If the matrix $C(1)$ has full rank *n*, then (6.9) shows that the permanent components of y_t can be seen as the sum of *n* independent random walks. On the other hand if the rank of $C(1)$ is $k < n$ then we can write:

$$C(1) = \gamma\delta'$$

where γ and δ are both $n \times k$ matrices of rank *k*. Consequently (6.8) can be written as:

(6.11) $$y_t = \gamma\left(\delta'\sum_{j=1}^{t} e_j\right) + d_t = \gamma\tau_t + d_t$$

where the $k \times 1$ vector $\tau_t = \delta' \Sigma_{j=1}^{t} e_j$ contains the k "common trends" of y_t. It follows from (6.11) that the common trends τ_t are random walks, since:

$$\tau_t = \tau_{t-1} + \delta' e_t$$

so equation (6.11) says that when the rank of $C(1)$ is $k < n$, the trend components of y_t can be viewed as the sum of *k* (instead of *n*) independent random walks τ_t. This result is known as "common trend representation" and was derived by Stock and Watson (1988).

Furthermore the fact that $C(1)$ has rank $k < n$, implies that there will be $r = n - k$ linear combinations of y_t, say $\alpha' y_t$, which are stationary, that is there are *r cointegrating combinations* for the variables y_t. The *r* independent *cointegrating vectors*, which are stacked into the columns of the $n \times r$ matrix α, are all orthogonal to $C(1)$, so:

(6.12) $$\alpha' C(1) = 0$$

This result, together with (6.11), gives:

(6.13) $$\alpha' y_t = \alpha' d_t$$

so the stationary linear combinations of y_t depend only on the transitory components d_t and not on the common trends τ_t. The result expressed by (6.13) suggests to Vahid and Engle (1993) the following question: are there linear combinations $\tilde{\alpha}' y_t$ of y_t which only depend on the common trends τ_t? This would imply:

(6.14) $$\tilde{\alpha}' d_t = 0$$

where $\tilde{\alpha}$ is a $n \times s$ matrix. Those authors show that (6.14) holds if and only if $\tilde{\alpha}$ satisfies the following relation:

(6.15) $$\tilde{\alpha}' \Delta y_t = \tilde{\alpha}' e_t$$

that is if and only if there exist *s* linear combinations of Δy_t which are an innovation with respect to all the information available at time $t-1$. If (6.15) holds we say that Δy_t have a *serial correlation common feature*. Every column vector of $\tilde{\alpha}$ is called a *cofeature vector* and defines a *cofeature combination*.

From the definition of the transitory component d_t in the multivariate Beveridge and Nelson decomposition, equation (6.14) implies:

(6.16) $$\tilde{\alpha}' C_j^* = 0 \quad \text{for every } j = 0,1,...$$

Vahid and Engle (1993) show that a consequence of (6.16) is that every C_j^* can be factorised as follows:

(6.17) $$C_j^* = F \tilde{C}_j$$

where F and $\tilde{C}_j$ are, respectively, $n \times (n-s)$ and $(n-s) \times n$ matrices. The matrix F is the same for all j and has rank $(n-s)$, while the matrices $\tilde{C}_j$ have rank less than or equal to $(n-s)$. Substituting (6.17) into (6.10) leads to the following expression for the transitory component:

$$d_t = F \sum_{j=0}^{\infty} \tilde{C}_j L^j e_t = F\tilde{d}_t \tag{6.18}$$

So the n-dimensional transitory component d_t can be viewed as a linear combination of $(n-s)$ "common cycles" $\tilde{d}_t$. Substituting (6.18) into (6.11) gives the so called "common trend - common cycles" decomposition for y_t that is:

$$y_t = \gamma\tau_t + F\tilde{d}_t \tag{6.19}$$

One corollary of (6.15) is that:

$$\tilde{\alpha}' C_0^* = \tilde{\alpha}'(I_n - C(1)) = 0 \tag{6.20}$$

Equation (6.20) characterises the matrix $\tilde{\alpha}$ with respect to $C(1)$: it follows, in fact, that $C(1)$ must have s eigenvalues equal to one, and that $\tilde{\alpha}$ is the matrix of the corresponding eigenvectors. Recalling, from (6.12), that the cointegrating vectors, stacked in the matrix α, are the eigenvectors corresponding to the r eigenvalues of $C(1)$ equal to zero, it follows that the columns of $\tilde{\alpha}$ are linearly independent from the columns of α.

Furthermore this implies that if n $I(1)$ variables y_t are linked by $r < n$ cointegrating combinations, so that they have $n-r$ common trends, they can have at most $n-r$ cofeature combinations, i.e. at most r common cycles. In other words, the sum of the number of common trends and the number common cycles cannot exceed n.

In the special case in which there are r cointegrating combinations and exactly $n-r$ cofeature combinations, equation (6.19) yields a unique decomposition of y_t into permanent and transitory components. In fact consider the $n \times n$ matrix $\boldsymbol{A}$ obtained by:

$$\boldsymbol{A} = \begin{bmatrix} \tilde{\alpha}' \\ \alpha' \end{bmatrix}$$

Due to the linear independence of α and $\tilde{\alpha}$, $\boldsymbol{A}$ will have full rank and will have an inverse $\boldsymbol{A}^{-1}$. This inverse can be partitioned as:

$$\boldsymbol{A}^{-1} = [\tilde{\alpha}^- \quad \alpha^-]$$

where $\tilde{\alpha}^-$ and α^- are, respectively, $n \times (n-r)$ and $n \times r$ matrices. Equations (6.13) and (6.14) imply that:

$$\boldsymbol{A} y_t = \begin{bmatrix} \tilde{\alpha}'\gamma\tau_t \\ \alpha' F\tilde{d}_t \end{bmatrix}$$

so that the series y_t can be partitioned as:

(6.21) $$y_t = \boldsymbol{A}^{-1}\boldsymbol{A}y_t = \tilde{\alpha}^{-}\tilde{\alpha}'\gamma\tau_t + \alpha^{-}\alpha' F\tilde{d}_t = \tilde{\alpha}^{-}\tilde{\alpha}' y_t + \alpha^{-}\alpha' y_t$$

It is possible to show that the matrices $P = \tilde{\alpha}^{-}\tilde{\alpha}'$ and $I - P = \alpha^{-}\alpha'$ in (6.21) are idempotent and so they can be viewed as projection operators. So if the number of common trends and the number of common cycles sum exactly to *n*, it is possible to determine a projection operator which, for every *t*, divides the innovation e_t into the innovation in the permanent component and innovation in the transitory one.

6.4 The multivariate unobserved component decomposition of Harvey

The model proposed by Harvey and discussed in section 4.6 can be easily generalised to the multivariate case where more variables are involved. Let y_t represent a column vector containing *n* variables y_{it}, $i = 1,...,n$. Each y_{it} allows for a decomposition scheme $y_{it} = g_{it} + c_{it} + \varepsilon_{it}$ in which the permanent and the transitory components can be modelled as in section 4.6.

The main point of this multivariate decomposition is to stress if and how the stochastic components on which each variable y_{it} is decomposed depend on several common factors.

For the growth variable g_{it}, the presence of common factors corresponds to the existence of cointegrating relations among the variables y_{it}. Concerning the cyclical component the presence of common factors can be viewed as an indication of similarities or synchronisation among the components c_{it} of the variables y_{it}. Similar cycles are characterised by the same structural characteristic (same periodicity and reduction factor) but they are not synchronised since they are not generated by the same sequence of innovations. So resulting cycles can be out of phase and of different amplitude. Common cycles are not only similar but they are also generated by innovations which are perfectly correlated. They can differ only in terms of their amplitude.

In order to obtain a parsimonious model for the variables y_{it} it is useful to reduce the number of innovations determining the *n* cyclical movements c_{it}. This can be possible by identifying $k < n$ common innovation factors which can be assumed as generators of the cyclical movements of our variables.

When $k = 1$ we are in a very special case when all *n* variables y_{it} are characterised by common cycles which are perfectly synchronised. When $1 < k < n$ the cyclical movements of each y_{it} is defined by a specific weighted combination of the *k* common innovation factors.

The reduction of *k* can be really useful to improve and simplify the inference of the estimated model. As an example we fix $n = 3$ so that $y_t = (y_{1t} \quad y_{2t} \quad y_{3t})'$.

In this case, if $k = 1$ each cycle c_{it} will be a linear transformation of the common elementary cycle c_t^{ω}, which is generated by the following equation:

$$\begin{pmatrix} c_t^{\omega} \\ c_t^{\omega*} \end{pmatrix} = \begin{pmatrix} 1-\rho L\cos\lambda & \rho L\sin\lambda \\ -\rho L\sin\lambda & 1-\rho L\cos\lambda \end{pmatrix}^{-1} \begin{pmatrix} \kappa_t \\ \kappa_t^{*} \end{pmatrix}.$$

Harvey proposed to write the model in the following way:

$$\begin{aligned} y_{1t} &= g_{1t} + c_t^{\omega} + \varepsilon_{1t} \\ y_{2t} &= g_{2t} + \theta_2 c_t^{\omega} + \varepsilon_{2t} \\ y_{3t} &= g_{3t} + \theta_3 c_t^{\omega} + \varepsilon_{3t} \end{aligned}$$

The only difference among the three cycles is represented by its amplitude and they are proportional to each other. When $k = 2$ each cycle c_{it} will result from a combination of two elementary cycles c_{1t}^{ω} and c_{2t}^{ω}. The two elementary cycles are similar but defined by different and independent innovation chronology. Our model can now be written in the following form:

$$\begin{aligned} y_{1t} &= g_{1t} + c_{1t}^{\omega} + \varepsilon_{1t} \\ y_{2t} &= g_{2t} + \theta_{21} c_{1t}^{\omega} + c_{2t}^{\omega} + \varepsilon_{2t} \\ y_{3t} &= g_{3t} + \theta_{31} c_{1t}^{\omega} + \theta_{32} c_{2t}^{\omega} + \varepsilon_{3t} \end{aligned}$$

The synchronisation of the cycles c_{it} will be only partial and it will depend on the parameters θ_{21}, θ_{31} and θ_{32}. The Harvey representation is very useful but nevertheless conventional. The cyclical component c_{1t} corresponding to the observed variable y_{1t} is assumed to be equal to the first elementary cycle c_{1t}^{ω}. The cyclical component c_{2t} is the sum of a multiple of the first elementary cycle and the second elementary cycle: $c_{2t} = \theta_{21} c_{1t}^{\omega} + c_{2t}^{\omega}$. The cyclical component c_{3t} is a linear combination of the two elementary cycles: $c_{3t} = \theta_{31} c_{1t}^{\omega} + \theta_{32} c_{2t}^{\omega}$.

When k common factors determine the cycles of n variables the formulation presented above is not the only one admissible. It appears clearly that the optimal possible identification of k common factors generating the n, innovations κ_{it}, $i = 1,...,n$, determining the cyclical movements c_{it} is a crucial point of this decomposition. The number of common factors is given by $rank(\sigma^2\Omega)=\sum^{n}_{j=1}\sum^{T}_{t=1}k_{j,t}^{2}$ which is the rank of the variance covariance matrix of innovations. When $rank(\sigma^2\Omega) = n$ it is impossible to identify a subset of factors generating the n cycles. In this case, even if the cycles can appear to be correlated, the phase shift among them is important and complex enough to avoid the possibility of identifying common factors.

6.5 The Long-run Restrictions decomposition

The long run restriction models are essentially based on the theory developed by Blanchard and Quah (1989). The authors use structural VAR models and they impose long run restrictions to obtain an estimation of potential output and the output gap conditional on such restrictions. The model proposed by Blanchard and Quah is a simple supply demand model where the long-run restrictions are represented by the fact that nominal shocks do not have any permanent effect on the output variable. Lalonde *et al.* (1998) and Funke (1997) have proposed extensions and applications of this approach.

Let z_t be a vector of *n* stationary variables. The Wold theorem allows this vector to be expressed in the following form:

$$z_t = \delta_t + C(L)\varepsilon_t \tag{6.22}$$

where δ_t is deterministic, $C(L) = \Sigma_{j=0}^{\infty} C_j L^j$ is the moving average matrix with $C_0 = I_n$ by definition. ε_t is a *n*-dimensional vector of one-step forecasting errors of z_t given z_{t-1} with $\varepsilon_t \sim N(0,\Omega)$. The long-run restriction approach assumes that the vector z_t has the following structural representation:

$$z_t = \delta_t + \Gamma(L)\eta_t \tag{6.23}$$

where $\Gamma(L) = \Sigma_{j=0}^{\infty} \Gamma_j L^j$ and η_t is an *n*-dimensional vector of structural shocks that are assumed standardised, that is $\eta_t \sim N(0, I_n)$. It is possible to identify the structural form (6.23) from the reduced form (6.22) by imposing the following restrictions.

$$\begin{gathered} \Gamma_0 \Gamma_0^{'} = \Omega \\ \varepsilon_t = \Gamma_0 \eta_t \\ C(L) = \Gamma(L)\Gamma_0^{-1} \end{gathered} \tag{6.24}$$

From equations (6.24) we obtain

$$C(1)\Omega C(1)' = \Gamma(1)\Gamma(1)'$$

This relation means that we can identify the matrix Γ_0 by imposing a sufficient number of constraints on the matrix $\Gamma(1)$ that describes the long-term relationships.

In the Blanchard and Quah (1989) approach it is assumed that the first variable of the vector z_t is production and that $\Gamma(1)$ is upper-triangular so that certain shocks are constrained not to have any long term effects on production. In particular, Blanchard and Quah make the hypothesis that nominal shocks do not have long-run effects on real variables.

Supposing, as in Blanchard and Quah (1989), that the logarithm of production is the first variable of the vector z_t and according to the long-run restrictions approach, we can obtain the following decomposition:

$$z_{1t} = \mu + \Gamma^g(1)\eta_t^g + \Gamma^{g^*}(L)\eta_t^g + \Gamma^d(L)\eta_t^d$$

where $z_{1t} = \Delta y_t$ is the first logged difference of real output, η_t^g is the vector of permanent shocks affecting the production, η_t^d is the vector of transitory shocks, $\Gamma_{z1}^g(1)$ is the long-run multiplier of permanent shocks and $\Gamma^{g^*}(L) = \Gamma^g(L) - \Gamma^g(1)$ describes their transitory dynamics.

The estimated potential output based on this approach is then given by:

$$g_t = \mu + \Gamma^g(1)\eta_t^g + \Gamma^{g*}(L)\eta_t^g$$

such that potential output is defined as the permanent part of production. The transitory component of production is given by $d_t = \Gamma^d(L)\eta_t^d$ and corresponds to the output gap.

A simple bivariate model provides an illustration of this approach. We suppose that z_t contains the logarithmic growth rate of production and inflation growth rate. The basic assumption is that inflation does not have permanent effects on production (see Chagny and Döpke, 2001). In other word, in the Blanchard and Quah model, it is assumed that the nominal variables are neutral with respect to the real ones.

In a more complex situation some transitory shocks may affect the real output but not inflation. Consequently, these shocks do not affect the output gap. It is then possible to identify some shocks that affect permanently inflation but not production. In the same way we can identify shocks that do not have any permanent effects on both output and inflation.

An extension of the model of Blanchard and Quah (1989) has been discussed by Lalonde et al. (1998). In their approach the SVAR contains more than two variables and the inflation is the second variable in the vector z_t. Once more, $\Gamma(1)$ should be upper triangular to reflect the long-run restrictions.

Production is then decomposed as follows:

$$z_{1t} = \mu \ + \Gamma^g(1)\eta_t^g + \Gamma^{g*}(L)\eta_t^g + \Gamma^{dp}(L)\eta_t^{dp} + \Gamma^{dd}(L)\eta_t^{dd}$$

where η_t^{dp} is the vector of shocks having transitory effects on production but permanent ones on inflation and η_t^{dd} is the vector of shocks that do not have any permanent effects on both production and inflation. Matrices $\Gamma^{dp}(L)$ and $\Gamma^{dd}(L)$ are defined according to the vectors η_t^{dp} and η_t^{dd}, and represent the multipliers of such shocks. The dynamics $\Gamma^{dp}(L)\eta_t^{dp}$ can be then used as an estimation of the output gap.

This approach is commonly referred to as long-term restriction on production and inflation. It gives a more constrained estimation than the long term restriction on production method proposed by Blanchard and Quah.

By defining:

$$\Gamma_{z1}^d(L)\eta_t^d = \Gamma_{z1}^{dp}(L)\eta_t^{dp} + \Gamma_{z1}^{dd}(L)\eta_t^{dd}$$

we see that the estimated output gap satisfies:

$$g_t = \Gamma^{dp}(L) = \Gamma_{z1}^d(L)\eta_t^d - \Gamma_{z1}^{dd}(L)\eta_t^{dd}$$

So the output gap corresponds to the transitory component of production, excluding those shocks which are not permanently related to inflation.

For the long-run restrictions on production we assume that the real output is integrated of order one. In the long-run restriction on production and inflation we assume that both production and inflation are $I(1)$. An interesting application of this extended version of the Blanchard and Quah model is proposed by Lalonde et al. (1998) to the U.S., Canada and German economies.

References

Adams C., P. Fenton, F. Larsen (1987), "Potential Output in Major Industrial Countries", IMF *Staff Studies for the World Economic Outlook.*

Astolfi R., D. Ladiray and G.L. Mazzi (2001), "Business Cycle Extraction of Euro-zone GDP: Direct versus Indirect Approach", *DIW Vierteljahrsheft*, 3-2001, 377-397.

Baxter M. and R. King (1995), "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series", NBER *Working Papers*, Cambridge.

Beveridge S. and C.R. Nelson (1981), "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with particular attention to the Measurement of the Business Cycle", *Journal of Monetary Economics*, 7, 151-174.

Blanchard O. and D. Quah (1989), "The Dynamic Effect of Aggregate Demand and Supply Disturbances", *American Economic Review*, 79, 655-174.

Boschan C. and G. Bry (1971), "Programmed Selection of Cyclical Turning Points", *Cyclial Analysis of Time Series: Selected Procedures and Computer Programmes*, NBER.

Boschan C. and W.W. Ebank (1978), "The Phase Average Trend: a new way of measuring Economic Growth", *Proceedings of the Business and Economic Statistics Section of the American Statistical Association*, Cowels Foundation Paper, 332, Yale University.

Boschen J. and L. Mills (1990), "Monetary Policy with a New View of Potential GNP", *Federal Reserve Bank of Philadelphia Business Review*, July-August, 3-10.

Burns A.F. and W.C. Mitchell (1946), "Measuring Business Cycles", NBER *Working Papers*, Cambridge.

Canova F. (1993), "Detrending and Business Cycles Facts", CEPR, *Discussion Paper*, 782.

Chagny O. and J. Döpke (2001), "Measures of the Output Gap in the Euro-Zone: An Empirical Assessment of Selected Methods", *Kiel Working Paper*, 1053, Kiel Institute of World Economics.

Cogley T. and Nason J.M. (1995), "Effects of the Hodrick - Prescott filter on Trend and Difference Stationary Time Series, Implications for Business Cycle Research", *Journal of Economic Dynamics and Control*, 19, 253-278.

Conway P. and B. Hunt (1997), “Estimating Potential Output: A Semi-Structural Approach”, Reserve Bank of New Zealand, *Discussion Paper*, G97/9.

de Brouwer G. and N. Ericsson (1995), “Modelling Inflation in Australia”, *Research Discussion Paper*, 9510, Reserve Bank of Australia.

de Brouwer G. (1998), “Estimating Output Gaps”, *Research Discussion Paper*, 9809, Reserve Bank of Australia.

European Commission (2000), “European Economy”, Brussels.

Evans G. and L. Reichlin (1994), “Information, Forecasts and Measurement of the Business Cycle” *Journal of Monetary Economics*, 33, 233-254.

Friedman M. (1957), “A Theory of the Consumption Function”, Princeton University Press.

Frisch R.A.K. (1933), “Propagation Problems and Impulse Problems” *Economic Essays in Honour of Gustav Cassel*, London, Allen & Unwin, reprinted in “Readings in Business Cycles”, ed. R.A. Gordon and L.R. Klein, London, Allen & Unwin 1966.

Funke M. (1997), “Supply Potential and Output gaps in West German Manufacturing”, *International Journal of Forecasting*, 13, 211-222.

Guay A. and P. St-Amant (1997), “Do the Hodrick-Prescott and Baxter-King Filters Provide a Good Approximation of Business Cycles?”, WP Center for Research on Economic Fluctuations and Employment (CREFE), 53.

Harvey A.C. (1985), “Trends and Cycles in Macroeconomic Time Series” *Journal of Business and Economic Statistics*, 3, 216-230.

Harvey A.C. (1989), “Forecasting, Structural Time Series Models and the Kalman Filter”, Cambridge University Press.

Harvey, A. C. and A. Jaeger (1993), “Detrending, Stylized Facts and the Business-Cycle.”, *Journal of Applied Econometrics*, 8, 231-47.

Henderson R. (1916), “Note on Graduation by Adjusted Average”, *Transactions of the Actuarial Society of America*, 17, 43-48.

Henderson R. (1924), “A New Method of Graduation”, *Transactions of the Actuarial Society of America*, 25, 29-40.

Hodrick R.J. and E.C. Prescott (1980), “Post-war US Business Cycles: an Empirical Investigation”, *Discussion Paper*, 451, Carnegie-Mellon University.

Hodrick R.J., and E.C. Prescott (1997), “Post-war U.S. Business Cycles: An Empirical Investigation”, *Journal of Money, Credit, and Banking*, 29, 1-16.

King R.G. and S.T. Rebelo (1993), “Low Frequency Filtering and Real Business Cycles”, *Journal of Economic Dynamics and Control*, n°17(1-2), 207-231.

Kydland F. and E. Prescott (1990), “Business Cycles: Real Facts and Monetary Myths”, Federal Reserve Bank of Minneapolis, *Quarterly Review*, Spring.

Lalonde R., J. Page and P. St-Amant (1998), “Une nouvelle méthode d’estimation de l’ecart de production et son application aux États-Unis, au Canada et à l’Allemagne”, *Working paper*, 98-21, Bank of Canada.

Laxton D. and R. Tetlow (1992), “A Simple Multivariate Filter for the Measurement of Potential Output”, *Technical Report*, 59, Bank of Canada.

Le Bihan H., H. Sterdyniak, P. Cour (1997), "La notion de croissance potentitelle a-t-elle un sens?", CEPII *Economie Internationale*, 69.

Lucas R. E. (1976), "Econometric Policy Evaluation: A Critique", *Carnegie Rochester Conference Series on Public Policy*, 1, 19-46.

Lucas R. E. (1980), "Methods and Problems in Business Cycle Theory", *Journal of Money, Credit and Banking*, 21, 696-715.

Nelson C. and H. Kang (1991), "Spurious Periodicity in Appropriately Detrended Time Series", *Econometrica*, 741.

Nelson C. and C.I. Plosser (1982), "Trends and Random Walks in Macroeconomic series", *Journal of Monetary Economics*, 10, 139-162.

Okun A.M. (1962), "Potential GNP: its Measurement and Significance", *Proceedings of the Business and Economic Statistics Section of the American Statistical Association*, Cowels Foundation Paper, 190, Yale University, reprinted in "The Political Economy of Prosperity", A.M. Okun, The Brookings Institution, Washington DC 1970.

Phillips A. (1958), "The Relationship between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957", *Economica*, 25(100), 283–299.

Sefton J.A. (1996), "Output Gaps and Inflation: An Approach using a Multivariate Beveridge and Nelson Decomposition", NIESR, mimeo.

Singleton K. (1988), "Econometric Issues in the Analysis of Equilibrium Business-Cycle Models", *Journal of Monetary Economics*, 21, 361-86.

Stock J.H. and M.W. Watson (1988), "Testing for Common Trends", *Journal of the American Statistical Association*, 83, 1097-1107.

Turnovsky S.J. (1995), "Methods of Macroeconomic Dynamics", MIT Press, Cambridge MA.

Vahid F. and R.F. Engle (1993), "Common Trends and Common Cycles", *Journal of Applied Econometrics*, 8, 341-360.

Watson M.W. (1986), "Univariate Detrending Methods with Stochastic Trends", *Journal of Monetary Economics*, 18, 49-75.

Whittle P. (1963), "Production and Regulation", Van Nostrand, Princeton, NJ.