Università degli Studi di Bergamo
Dottorato di Ricerca in
Metodi Computazionali per le Previsioni
e Decisioni Economiche e Finanziarie

XXIV Ciclo

Estimating Business Cycle: from Bandpass Filters to Eurocoin Approach

Tesi di Dottorato di: Antonio Frenda

Settore Scientifico Disciplinare della Tesi: SECS-P/05 Econometria

Supervisori: Prof. Marco Lippi
Prof. Giovanni Urga

Author's e-mail address: frenda@istat.it; frendaantonio@libero.it

Acknowledgments

I thank the Supervisors Prof. Marco Lippi (University La Sapienza-Rome) and Giovanni Urga (University of Bergamo and Cass Business School) for the thoughtful supervision of this thesis.

A special thank to Prof. Marida Bertocchi (Director of the Doctoral Program) and Valeria Caviezel (University of Bergamo) for their steady encouragement during last years and for insightful advices that improved the quality of this work.

I am grateful to Lorenzo Bencivelli (Bank of Italy) for helpful explications concerning Eurocoin Indicator production.

This thesis also benefits of comments from the participants to the Interdisciplinary Workshop on Econometric and Statistical Modelling of Multivariate Time Series (Univ. Catholique de Louvain-la-Neuve, 25-27 May 2011) and to the PhD Student Research Day (University of Bergamo, 18 October and 17 December 2010).

Dedicated to my beloved daughters Julia and Anita and to my wife Jadwiga.

Contents

Introduction

The analysis of large-dimensional dataset plays an essential role in econometric research, particularly in the field of economic forecasting. The most important indicator of economic activity is the GDP (Gross Domestic Product) and, unlike the industrial production (IP), it comprises services, agriculture, and the public sector. Unlike surveys, GDP does not contain any subjective assessment. And, different from IP, GDP is defined only quarterly and with a lag.

Monthly indicators are commonly used in the prediction of current data on GDP before the data are available. The main groups of indicators generally used are: Surveys, Financial market data; Labour market data; Monetary aggregates; Industrial production; Prices; Demand Indicators; Foreign Trade.

The Eurostat Handbook on Quarterly National Accounts (2000) defines a flash estimate:

"…the earliest picture of the economy according to national accounts concepts, which is produced and published as soon as possible, after the end of the quarter, using a more incomplete set of information than that used for traditional quarterly accounts."

For the euro area, a flash estimate of GDP is released by Eurostat about six weeks after the end of the reference quarter, and a full set of indicators for the second quarter of the year is not available any earlier than the flash estimate of GDP.

New Eurocoin is a synthetic and up-to-date statistics measure of the Euro-Area conjuncture. The target of NE is c_t, the medium to long-run component of the Euro Area GDP growth. It is a performance measure published monthly by the Bank of Italy and CEPR. Estimations of NE are obtained through the generalized dynamic factor model. The specific references are Altissimo and others (2001, 2009), Forni, Hallin, Lippi and Reichlin (2001, 2004, 2005).

In the dynamic factor model approach, a vector of “n” time series is decomposed into two mutually orthogonal components: a common component characterized by few common factors or latent shocks, and a component "idiosyncratic", led by n specific shocks (one for each variable in the panel). These models allow a net reduction of the cross-sectional size of the dataset.

Since New Eurocoin indicator provides a summary index of the medium to long-run component (MLRG) of the GDP solely for the whole Euro area aggregate, the main contribution of this research is to propose some procedures to estimate the following disaggregated components of European GDP:

- Sectoral smoothed growths;
- MLRG concerning some countries belonging to the Euro Area;
- Expenditure components.

The Eurocoin approach is therefore used in this thesis to estimate monthly smoothed GDP components for the following reasons:

1. Euro Area flash estimate is released on a quarterly frequency, with a certain delay and may be subject to significant revisions afterwards.

2. GDP growth in any quarter depends on seasonal effects and measurement errors; therefore GDP can be misleading;

3. GDP can be influenced by factors affecting only a particular country. These factors are not important for outlining the health of the euro area economy as a whole, but they can be used to asses the country's national cycle.

Furthermore, It is a well-known result in the literature isolating the business cycle in integrated series that band-pass filter could deteriorate at the end of the sample, being less reliable for the most recent data. This is very relevant for economic policy. Altissimo, Cristadoro, Forni, Lippi, Veronese (2008) show that the same problem arises with application to stationary time series. And, through New Eurocoin, they develop a method to obtain smoothing of a stationary time series so as to avoid the occurrence of end-of-sample deterioration and short-run fluctuations.

In synthesis, the primary objective of this thesis is to produce smoothed growth indicators that describe the behaviour of economic activity for a large number of countries and sectors at a monthly frequency, while utilizing a wide range of economic time series in a timely fashion. The innovation of this research are some procedures, based on Eurocoin methodology, to estimate MLRG concerning some countries belonging to the Euro Area; sectoral growth; smoothed GDP components (Consumptions, Investments, Foreign Trade).

In this thesis, some disaggregated versions of the Eurocoin indicator are proposed; it proceeds as follows.

In the *chapter 1* we present:

- a brief review of filtering methods and short-term forecasting approaches for GDP known in literature and commonly used by practitioners;
- the advantages of the Eurocoin approach and the criteria to choose the common factors in the generalized dynamic factor model are explained;
- classic and modern approaches for the measurement of business cycle and to detect the turning points in each cycle phase.

In *chapter 2* a new theoretical framework for disaggregated business cycle analysis is proposed. Data are described and in sample estimates are produced. Goals developed are the following:

a) using Eurocoin approach in building sectoral analysis of the medium to long-run component of GDP growth **(MLRG)** in the European economy. In fact, the 157 basic variables that now constitute the complete dataset from Thomson Financial Datastream, used to process Eurocoin indicator, belong to different homogeneous data groups.

Some series ignore large portions of economic activities (e.g industrial production and export), and all these series exibit heavy short-run fluctuations and could provide incoherent signals. Consequently, “there is much diversity and uncertainty about which indicator are to be used” (Zarnowitz and Ozyildirim, 2002). That is why, therefore, it is necessary to outline a more specific analysis of MLRG.

The strategy used has been the projection of sectoral added value on European factors (which are the combination of the 157 variables contained in the Thomson Financial Datastream, and used by the Bank of Italy to build Eurocoin).

b) we develop monthly smoothed growth indicators for the main Euro Area economies that can utilize a wide range of economic information in real time. We build these *national indicators*, by the projection of national or Euro Area GDP on European and national factors. This goal has been achieved for the following countries: Belgium, Italy, France, Germany, Spain. A national Eurocoin (NE) is the projection of bandpassed GDP on a set of regressors, which are linear combinations of national variables contained in the Thomson Financial Datastream used by the Bank of Italy (e.g. we will project German Gross Domestic Product on 39 German variables contained in the Dataset).

c) Improving Eurocoin methodology, dividing European variables used to build common factors in real and financial variables. Real variables, e.g. industrial production and foreign trade, concern real economic activities. This subdivisions among real economy and financial economy is substantially confirmed in Cristadoro, Forni, Reichlin, Veronese (2001). We show that a combination between “real MLRG” and “financial MLRG”, obtained projecting Euro Area GDP respectively on real and financial variables, can compete with the truncated band-pass filter within the sample, and it outperforms Eurocoin in terms of RMSFE. Since Newbold and Granger (1974), and more recently Kennedy (2001), there has been some agreement in literature that the ‘best forecasting method’ is a `combined forecast', formed as a weighted average of forecasts generated by a different technique. We show in chapter 1 that the minimum of variance error is less than the error variances of the individual forecasts;

d) Series contained in Thomson Financial Datastream “might provide contradictory signals”. Our analysis determines the contribution of different bandpassed medium to long-run components (consumption, investment, exports, imports) of GDP in the European business cycle. It is important, therefore, to analyze the volatility of Household Consumption, Investment, Imports and Exports, whether it is weak or strong, using Eurocoin methodology to calculate these components (actually Eurocoin is only used to calculate GDP cycle). Do these series appear to be consistently correlated? We can identify a different synchronization for the various components and a lead of the others.

Chapter 3 discusses real-time estimating exercises for the models investigated in chapter 2 and it presents the results. These estimates have been obtained simulating the situation one would have faced at the end of each month in terms of data availability.

The exercises that we develop use the estimates $\hat{c}_t(t+h)$, of each disaggregated indicators at time t using the data from 1 to t + h, h = 0, 1, 2, with t running from January 2003 to December 2010. Analysis of real time performance will regard:

- the ability of the real time indicator to approximate the target;
- the capacity of the indicators to signal the correct sign of the change;
- the size of the revision errors after one month;
- the ability to perform well in signalling turning points in the target.

The last section concludes and raises some questions for pursuing further research.

In synthesis, in this research we investigate about the relationships among European and national MLRG (aggregated, by economic sectors and GDP components), through the development of forecasting methodologies based on dynamic factor model and the combinations of real and financial variables contained in the Thomson Financial Datastream.

Chapter 1 - Review of the Literature

"There is…no need to qualify business cycle observations by restricting them to particular countries or time periods; they appear to be regularities common to all decentralised market economies".

R E Lucas (1977)

The innovation of this research is the development of disaggregated procedures to obtain smoothing of a stationary time series. The objective is to build sectoral and national real-time monthly estimates of GDP growth purified from seasonal and other short run fluctuations, as well as from errors in the measurement of GDP, which is highly reliable at the end of the sample. The application that we develop also eliminates local and sector-specific shock that are not useful to assess our objectives.

The main aim of this chapter is to outline the econometric literature useful for the analysis of the medium-long run components of growth. It is the theoretical reference on which chapter 2 and 3 of the thesis, with the empirical applications, are based.

Predominantly:

- in section 1.1, we show bandpass filters to be considered in this thesis a target for comparisons and a method to eliminate erratic components of the growth rate for different sectors and European countries. Also, classic filtering methods, not based on spectral analysis, will be outlined;
- in section 1.2, the two main approaches are presented for the short term estimation of Gross Domestic Product, i.e. *Dynamic factor models* (New Eurocoin is an application of this model) and *Bridge equations;*
- in section 1.3, the advantages of the Eurocoin approach and the criteria to choose the common factors in this model are explained;
- 1.4 is a short outlook to estimation methodologies that use the Kalman filter in the presence of missing observations at the end of the sample due to publication lags with regard to variables in the dataset. It seems useful to highlight the differences with the Eurocoin approach;
- In 1.5 we describe classic and modern approaches for the measurement of business cycle and to detect the turning points in each cycle phase;
- 1.6 and 1.7 regard, respectively, the combination of forecasts (on which the sections 2.4 and 3.4 of this thesis are based) and how to aggregate individual (national) time series to construct the series for the aggregate Euro Area.

As we show in this chapter, bandpass methods produce highly unreliable estimates at the end-of-sample. We will use generalized dynamic factor model on which Eurocoin Indicator (Altissimo et al., 2009) is based, published monthly by the Bank of Italy and CEPR, to develop our disaggregated indicators. In fact, in the academic literature models for short-term nowcasting can be subdivided into two classes: dynamic factor models and Bridge equations estimation.

Estimates that we develop are considered as disaggregated because they do not make reference to the entire Euro Area gross domestic product, contrary to the Eurocoin.
We focus on the medium to long-run component of the growth (MLRG), i.e. the smoothed component of GDP growth rate obtained by removing the fluctuations of period shorter than or equal to one year. MLRG bears no relationship to any specific definition of trend.

1.1 Filtering Methods for the Analysis of Business Cycle

Altissimo et al. (2009) have developed New Eurocoin (NE) indicator (to be outlined in paragraph 1.3), based on a generalized dynamic factor model, to smooth stationary series by using only contemporaneous values of a large dataset, so that no end-of-sample deterioration occurs.
The removal of short-term dynamics of stationary series from the medium to long-run growth can be obtained through band-pass filters: they are infinite moving averages based on past and future values of GDP, and they can deteriorate at the end of the sample, when the missing values useful to describe the cycle is replaced through predictions. These filters work well in the middle of the sample, but they perform badly at the beginning and end of the sample at the end of the sample.
Since estimation of the last data point in the bandpassed target is bad, we decide in the thesis to develop some disaggregated indicators by sectors, countries and expenditure components based on the Eurocoin approach.
These indicators are built in chapter 2, while in chapter 3 their real time performance is shown with respect to a measure of the "trend-cycle GDP growth" obtained in the middle of the sample by a bandpass bilateral filter on GDP growth components (see equation (3.1)).
The *cycle*[1] may be decomposed into proper cycle (medium to long-run growth) and irregular components, and there is some subjectivity in the definition of the business cycle (see Canova and Paustian, 2007).
The decomposition between trend, cycle and irregular components is not unique, and any method requires arbitrary choices. “A cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle. This sequence of changes is recurrent but non periodic. In duration, business cycles vary from more than one year to ten or twelve years. They are not divisible into shorter cycles of similar characters with amplitudes approximating their own.” (Burns and Mitchell, 1946, p.3.)
It is essential to separate the more recent literature that regards bandpass filters of Christiano-Fitzgerald and Baxter and King, based on frequency information that define the medium to long

[1] According to the traditional definition of the business cycle given by an expansion phase followed by a recession. The transition from one phase to another is determined by a contraction or a recovery (classic cycle). By "modern" definition (Lucas, 1977), the cycle is determined by fluctuations around the trend and regularity in the movement of aggregate economic series following an external disturbance (shock): deviation cycle.

run component of growth, and the classic methodology (Quah, Beveridge and Nelson) looking at the trend as a non-stationary component correlated with the cycle.

In this work we will compare the medium to long-run component of the growth with spectral bandpass filters. It is however worthwhile to describe other types so as to highlight differences and eventually some common aspects.

Particularly, different filtering methods are explained in literature either to remove short-term dynamics of stationary series or for analyzing the medium to long-run growth (MLRG). The following will be used in our research:

1) Weighted moving averages
2) Filters based on spectral representation
3) Methods based on generalized dynamic factor models

1.1.1 Some References to Classic Filters

A simple filter is based on the method of the "*first differences*" which assumes that:

$\Delta y_t = y_t^c$ with $\Delta y_t = y_t - y_{t-1}$

$y_t^T = y_{t-1}$

where y_t^c is the cycle and y_t^T the trend.

The medium to long term component of the growth purified by short-term volatility and statistical noise is obviously not observable, but, when we have a sufficient number of futures observations, it can be estimated with a two-sided bandpass filter: it can be defined in simple terms as a weighted moving average centered growth rates of quarterly GDP[2].

Estimating the cyclical component with quarterly data, it is possible to apply centered moving averages of order 5, that eliminate cyclical fluctuations equal or lower than 5 quarterly (Henderson, R. 1916)

$$y_t^{**} = \frac{1}{8}(y_{t-2} + 2y_{t-1} + 2y_t + 2y_{t+1} + y_{t+2})$$

The centered and weighted moving averages could also be approximated by a trend growth rate, i.e. a rough moving average that is not centered, with four terms.

However, we ought to pay attention to the fact that we can create spurious correlation effects (Slutsky-Yule effect) introducing spurious cyclical trends.

Another method based on simple decomposition between trend and cycle is the Beveridge and Nelson method (1981). It separates trend and cycle with a strong negative correlation. This decomposition is relatively consistent with the idea of the typical business cycle, and defines the

---

[2] A reference is Cristadoro et al. (2008), "L'indice punta alla recessione", *www.lavoce.info:* "Because the lack of future data at the end of the sample justifies the €-coin estimate, which provides a good approximation in real time".

trend as a random walk. So, this filter is based on the assumption of a particular relationship between cycle and trend (negative correlation between changes in trends and changes that generate the cycle).

OECD estimates, in various publications, potential GDP[3] by the *split time-trend method*: this approach is based on the identification of a linear trend for each of the different economic cycles that compose the estimated period, and the cycle is defined as the interval between two consecutive peaks:

$$\ln Y_t = \alpha_0 + \sum_{j=1}^{n} \alpha_j T_j + e_t$$

In each cycle a constant growth rate of potential output is therefore identified, and output gap estimates are located symmetrically around it. This type of method requires the arbitrary and preliminary definition of peaks, and it does not detect structural changes within a single cycle. It is also less useful in estimating current times unless we take strong assumptions about what will be the next peak.

1.1.2 Filtering Methods and the Spectral Representation

The modern spectral analysis using Fourier series is based on the assumption that the amplitude of sines and cosines are random variables. Filtering methods based on spectral representation formulate the smoothing and de-trending problems in the frequency domain, and approximate the ideal infinite band-pass filter.

Fourier (1822) showed that any periodic function can be expressed as a linear combination of an infinite number of sinusoidal functions[4] with different frequencies.

The first works that recognized the attraction of econometric analysis in the frequency domain are: Granger and Morgenstern (1963), Granger and Hatanaka (1964), Granger (1966), Sargent (1973).

These methods define the business cycle according to a specified range of periodicities.

A central result in time-series is that any stationary time series can be regarded as the sum of orthogonal sinusoidal components. The spectral representation of a stationary stochastic process is its decomposition into a sum of unrelated periodic functions[5].

---

[3] It is the level of output that an economy can produce at a constant inflation rate.

[4] The most simple sinusoidal function with constant amplitude is f(t) = $\rho \cos\left[2\pi(\omega t - \theta)\right]$. Representing the business cycle by sinusoidal functions, parameters concerning the function must be defined: the amplitude ρ of oscillations, the frequency ω (the number or fraction of oscillations, or cycles, per unit of time considered) and the phase θ, namely ' abscissa of the maximum point. The inverse of the frequency expresses the period of oscillation, the period. If the period is equal to 8 years for example, the frequency will be equal to 1/8 year cycle.

[5] The Cramer-Kolmogorov theorem on the spectral representation of a stationary stochastic processus $\{X_t\}$ allows to understand the influence of cycles of different frequencies on the behaviour of $\{X_t\}$.

For example,

$$Y_t = \int_0^{\pi} \cos \omega t du(\omega) + \int_0^{\pi} \sin \omega t dv(\omega)$$

where $\{Y_t\}_{t=0}^{\infty}$ represents a stationary stochastic process in discrete time, with $u(\omega)$ and $v(\omega)$ orthogonal processes defined on the open interval $(0,\pi)$. Under some weak additional assumptions, the existence of the band-pass filter is implicit in (1), because it implies that it is possible to decompose the stationary time series into components indexed by frequency ω.

Baxter and King and *Christiano-Fitzgerald* filters are based on spectral representation. Using filters to determine the cycle, it should be noted, according to Baxter and King (starting from the work of Burns and Mitchell,1946), that many U.S. economic series admit a cycle period between 6 and 32 quarters. Consequently, the trend, reflecting the evolution of long-term or low frequency, has a longer period of 32 quarters. The erratic component, which corresponds to changes in short-term or high frequency, has a period less than 6 quarters.

Baxter and King (1999) method is based on a priori choice of frequencies that define the cycle. This approach also isolates an irregular component: both the cycle and the erratic component extracted are stationary.

In the Baxter-King filter, the irregular (erratic) component s_t is supposed to correspond to high frequencies and our time series y_t used in this research to analyze MLRG have the following decomposition:

$$y_t = c_t + s_t = \beta(L) y_t + [1 - \beta(L)] y_t \tag{1.1}$$

s_t includes all the waves of period shorter than one year; c_t which is a smoothing of the GDP growth. It is the ideal target to which our possible estimates for medium to long-run growth are compared.

The filter of Baxter and King, outlined in Altissimo, Cristadoro, Forni, Lippi, Veronese (2009), refers to the representation of time series in the frequency domain based on the spectral representation theorem, which allows to decompose a process Z_t in a weighted sum of sinusoidal functions exp(itω) and the weight of each frequency is deducted from the spectral density of the process. The latter is the Fourier transform of the autocorrelogram process. The value of the spectral density of the process for the frequency ω is proportional to its contribution rate to the total variance of the process.

In the filter of Baxter and King represented below, the coefficients B_K are based on simple sinusoidal functions. It is a bilateral filter suitable as target for the control of performance concerning New Eurocoin and to identify medium to long-run component:

$$c_t = \sum_{k=-M}^{M} B_K \Delta GDP_{t-k} \tag{1.2}$$

The cycle, however, in Baxter and King, is defined as the stationary component of a series whose period is between 6 and 32 quarters. Isolate the cycle means then applying to the initial set the band-pass filter that preserves these frequencies and cancel the others. It is obtained classically as the difference between two low-pass filters. This filter must be symmetrical in order to not introduce phase shifts between the original series and the filtrate result. The ideal low-pass filter, associated with the frequency ω_0 must have a gain function[6] of the form identified by Brockwell and Davis, 1996 (chapter 4):

$$A(e^{-i\omega}) = \begin{cases} 1, if\ |\omega| \le \omega_0 \\ 0, if\ |\omega| > \omega_0 \end{cases}$$

The band pass filter for the frequency ω_0 is a linear transformation (infinite moving average) of a series that maintains some frequencies lower or equal to ω_0 and eliminates frequencies upper than ω_0. The band pass filter that is associated to the frequencies ω_1 and ω_2 (with $\omega_1<\omega_2$) maintains the frequencies included between ω_1 and ω_2 and eliminates the ones lower than ω_1 or upper than ω_2.

The series y^*_t created by the band pass filter can be expressed by the following moving average:

$$y_t^* = \sum_{-\infty}^{+\infty} a_k y_{t-k}$$

This means to apply the polynomial (L) to the series $y_t : a(L) = \sum_{-\infty}^{+\infty} a_k L^k$

Baxter and King have proposed a method for isolating cyclic fluctuations in economic applications. They show that the best approximation of this infinite filter with a finite order k symmetric filter consists of 2k +1 terms. This is obtained by the simple truncation of the infinite filter of the order k, provided that the sum of coefficients is equal to 1 adding the same correction to each coefficient. Of course, the accuracy of the finite filter increases with the number of elements, but it is compensated for with the loss of k points in each extreme of the series. In quarterly data, Baxter and King consider k=12, ie a symmetric filter with 25 terms, leading to the loss, at the end of the series, of 12 quarters. Baxter and King have proposed a finite approximation MA, that can be represented as follows:

$$y_t^T = \sum_{i=-k}^{k} \omega_i y_{t-i} \tag{1.3}$$

[6] The square root of the transfer function is called the gain function and it determines the variation of amplitude for a cyclical component following application of a filter.

where ω_i are weights derived by the inverse of the symmetric Fourier Transform of frequency response function.

Christiano-Fitzgerald (1999,2003) derived a filter, assuming that the data are generated by a pure random walk. They note that this assumption is not real for most macro series, but they argue that it produces a filter that works well in many circumstances.

The Christiano-Fitzgerald (CF) random walk filter is a band pass filter based on the same principles as the Baxter and King (BK) filter. The BK version is a symmetric approximation with no phase shifts in the filtered series.

So, the CF filter, which is an asymmetric version of the Baxter-King filter, can be calculated as follows:

$$c_t = B_0 y_t + B_1 y_{t+1} + \ldots + B_{T-1-t} y_{T-1} + \overline{B}_{T-t} y_t + B_1 y_{t-1} + \ldots. + B_{t-2} y_2 + \overline{B}_{t-1} y_1 \quad (1.4)$$

$$B_j = \frac{\sin(jb) - \sin(ja)}{\pi j}, j \geq 1, B_0 = \frac{b-a}{\pi}, a = \frac{2\pi}{p_u}, B = \frac{2\pi}{p_l} \quad (1.5)$$

$$\overline{B}_k = -\frac{1}{2} B_0 - \sum_{j=1}^{k-1} B_j \quad (1.6)$$

p_u and p_l represent the cut-off cycle length in the month. Cycles longer than parameter p_l and shorter than p_u are preserved in the cyclical component c_t.

Phase correctness and symmetry comes at the expense of series trimming. Depending on the trim factor, some values at the end of the series cannot be calculated[7].

The Christiano-Fitzgerald random walk filter uses the whole time series for the calculation of each filtered data point. An advantage is that it works well on a larger class of time series than the BK filter, and it converges, in the long run, to the optimal filter. In real time applications, CF can outperform BK.

In Chadha, Janssen, Nolan (2000), the following variables have been observed to analyze UK business cycle fluctuations since 1871: GDP Output, Consumption, Investment, Broad money, Real narrow, Wages, Prices, Interest rates, Current account balance. It appears that, comparing BK and CF filters, there isn't a large difference between each set of correlations and output, especially in terms of procyclicality[8] or countercyclicality. On the contrary, concerning ratio of variance of macroeconomic variables in pre and post World War II, results in BK and CF appear different, especially in the Investment variable.

[7] "There is a trade-off between the trimming factor and the precision with which the optimal filter can be approximated" *Oecd system of composite leading indicators*. November 2008.

[8] Procyclicality is the tendancy of financial variables to fluctuate around a trend during the economic cycle. Increased procyclicality means fluctuations with stronger amplitude following a shock. The path of asset prices and evolution of financial aggregates is broader with an irregular (sometimes non-linear) form of volatility.

Therefore, using different filters, it is possible to have some meaningful differences in relation to the volatility of macroeconomic time series. Nevertheless, the structure of correlations among filtered series generally turn out to be homogeneous.

A tool used in the past for the analysis of cycles is linear *Hodrick-Prescott* (HP) filter. It is used to obtain a smoothed non-linear representation as it is more sensitive for fluctuations in long term than in the short term. The assumptions on which the filter is built is that the series of macroeconomic variables y_t can be decomposed into trend and cyclical components and that the determinants of growth (the trend component) are subject only to slow and gradual changes:«Our prior knowledge is that the growth component varies “smoothly” over time» (p. 3)

Decomposing the series y_t into a cyclical component c_t and a trend component τ_t ,the parameters are obtained by solving the following optimization problem:

$$\min_{\tau_t} \sum_t (y_t - \tau_t)^2 + \lambda \sum_t (\tau_{t+1} - 2\tau_t + \tau_{t-1})^2$$

with $y_t = \tau_t + c_t$

The first term of the equation to minimize is the sum of squared deviations. The second is a multiple (with a factor λ) of the sum of squared second differences from the component trends. This second term tends to penalize changes in the rate of growth of the trend component. The higher is the value of λ, the higher the penalty. Hodrick and Prescott for quarterly data defining a value of 1600, as (“*…a 5% cyclical component is moderately large, as is a 1/8 of 1% change in the growth rate in a quarter…*”

The optimization problem has a solution that can be represented by a linear transform that is independent from Y_t as suggested Maravall and del Rio (2001). In the HP filter weighing parameter value is somewhat arbitrary and the choice of parameters is considered a critical point of the method of HP (Dolado et al., 1993). It could lead, in turn, to some relatively variable estimates quite close to the real data series (low values of λ) or not to variable estimates implying large deviations between real variable and estimated variables. No theoretical reason beyond conventions and common practice in literature leads to consider a value of λ more reliable than others (see Billmeier, 2004, p. 18-9).

For annual data, Baxter and King (1999) suggest the value λ=10, because it approximates a bandpass filter that removes from the cycle periodicities longer than 8 years. Dolado et al. (1993) employ λ = 400, dividing 1600 by 4 for quarterly and annual data, while Backus and Kehoe (1992), Giorno et al. (1995) as European Central Bank (2000) use a λ=100.

Dolado et al., for monthly data, recommend a default value λ = 4800, while the software E-views employs the default value 1440

1. 2 Short Term Forecasts of GDP Growth

The two main approaches have been used in empirical applications for short term forecast of Gross Domestic Product, also named *Nowcasting*[9], are:

- *Dynamic factor models* (New Eurocoin is an application of this model).

- *Bridge equations estimation*: they are regressions of quarterly GDP growth on a small group of key monthly indicators. This modelling strategy has been popular among political institutions that commonly combine various estimates of GDP from bridge equation models to examine information from a large number of predictors (see Baffigi, Golinelli, Parigi, 2004).

In this research, the former approach is considered; the *dynamic factor model* is designed to extract common movements from a large set of data series and to synthesise them into a few artificial latent factors, which represent the main sources of variation in the data set in order to nowcast the main smoothed components of European growth. However, the two approaches will be outlined in this section for further developments of the research.

In Banbura et al. (2010) *nowcasting* is defined as "the prediction of the present, the very near future and the very recent past", using timely monthly information to nowcast quarterly variables that are published with long delays[10].

The two methods aim to extract the underlying common tendencies across the set of indicators, as many indicators of economic activity tend to move closely together, from the idiosyncratic irregular movements concerning individual series.

Either dynamic factor models or bridge equations comprise the following two steps.

We will consider a forecasting equation that predicts quarterly GDP growth in a certain period *t* (gdp_t) from monthly indicator x_t:

$$gdp_t = c + \beta x_t + u_t \tag{1.7}$$

with c and u_t denoting a constant and a residual term respectively, while β denotes the coefficient related to the monthly indicator. Monthly observations for the indicator x_t are often incomplete within the quarter. The missing observations must be forecasted to obtain a quarterly value of x_t. Standard time series models may be used for this purpose, as is the case for the forecast system of bridge equations.

In a second step, as described in equation (1.7) above, GDP growth can then be predicted:

- in a dynamic factor model by the factors (linear combinations of observable monthly data);

9 As L. Reichlin explains in *The Eurocoin project: new modelling ideas at central banks* (CEPR – September 2007): "Should we care about the short-run? Yes! It is the only thing we can really forecast ".

10 They argue that "the nowcasting process goes beyond the simple production of an early estimate and it consists in the analysis of the link between the news in consecutive data releases and the resulting forecast revisions for the target variable".

- forecasting from bridge equations by *a* group of key monthly indicators: Diron (2006) employs 8 equations to forecast euro area GDP, making use of data concerning industrial production, construction output, retail sales and unemployment.

1.2.1 Estimation of GDP by Bridge Equations

This is a traditional method for obtaining an initial estimate of quarterly GDP growth using information on monthly variables that links monthly data, usually issued in the early months of the quarter, with quarterly data, such as GDP and its components released late and available with monthly frequency.

In this modelling strategy, we will denote a vector of k stationary monthly indicators $x_t^j = (x_{1,t}^j,....,x_{k,t}^j)'$. For every bridge equation j, $t = 1,....T$, forecasts of monthly predictors are typically based on univariate time series models. While the bridge equation is estimated from quarterly aggregates of monthly data.

Finally, the resulting values will be used as regressors in the bridge equation to obtain the GDP forecast:

$$y_t^Q = \mu + \sum_{i=1}^{k} \beta_i^j(L) x_{it}^{jQ} + \varepsilon_t^{jQ}$$

where μ is ad intercept parameter and $\beta_i^j(L) = \beta_{i0}^j + ...\beta_{is_i^j}^j L^{s_i^j}$ represents lag polynomials of length s_i^j.

Bridge equations should generally be cross-checked against each other because it may be misleading to rely only on one of them due to uncertainties in results. Extended models are based on equations to forecast the aggregate euro area expenditure components of GDP, from which another forecast for the euro area GDP can be derived. For each of the demand components, the forecast is obtained as the average of forecasts from about ten equations. There are many examples in literature, among others, showing the good results of bridge approach in terms of forecasting accuracy: Buffettau and Mora (2000), Baffigi, Golinelli, and Parigi (2002-2004), or more recently, Diron (2008).

The main implementations of this technique:

- an approach, implemented at the ECB, which combines a set of bridge equations selected based on multiple regressions, called *BES method* (Bridge Equations based on selected predictors), with a few equations using the selected indicators (see Diron, 2006; Rünstler and Sédillot, 2003);

- a bridge equation approach based on all predictors (*BEA models*), which combines the estimates of GDP, based on a large number of equations with only one predictor each (see

Kitchen and Monaco, 2003). The idea of model averaging to combine information from large data sets has been discussed extensively in Clements and Hendry (2004).

Figure 1.1 Bridge equations used to forecast Euro area GDP growth (BES model)

	Equation											
Explanatory variables	1	2	3	4	5	6	7	8	9	10	11	12
Industrial production (total)	*	*	*	*	*	*	*	*	*			
Ind production construction		*	*	*	*	*		*	*			
Retail sales		*	*	*	*	*	*	*	*			
New car registrations	*	*	*	*						*		
Service confidence				*	*		*		*			
Unemployment rate					*	*						
Money M1								*	*			
Business confidence										*		
EuroCoin (CEPR)											*	
OECD leading indicator												*

Source: Angelini, Mendez, Giannone, Reichlin, Rünstler (2008)

1.2.2 A brief History of the Dynamic Factor Model

Growth indicators can be estimated using the dynamic factor model (DFM). DFM is particularly useful because it utilizes a large number of economic time series in a timely fashion.

Through DFM approach, factors can be constructed from monthly data set and forecasted over the desired horizon. GDP growth can then be predicted from the factors in a second step. The factors are linear combinations of the observable monthly data.

Economic activity in market economies is characterized by cyclical behaviour and comovements[11] in macroeconomic variables. If co-movements are strong, the state of the economy could be represented by an index reference cycle that describes the common behaviour of such variables. This idea, first suggested by Burns and Mitchell (1946), is behind the NBER indicator. So, the formal model that best captures the reference cycle is the *dynamic-factor model* that was proposed by Sargent and Sims (1977) and Geweke (1977). A vector of n time series is represented as the sum of two non observable orthogonal components; a common component driven by few common factors, and an idiosyncratic component composed by n idiosyncratic factors. If only one common factor affects all of the time series contemporaneously (i.e., without lags), this factor can be considered as the reference cycle (Stock & Watson, 1989). Factor models can be used to learn about macroeconomic behaviour on the basis of disaggregated data (regions and sectors)[12].

[11] • Depending on the nature of comovements, macroeconomic variables are divided into: - pro-cyclical variables tending to go in the direction of the movements of GDP (eg consumption, investment, employment) - anti-cyclical variables, whose performance tends to be negatively correlated to that of GDP (eg, unemployment) - variables acyclic, that is little affected by fluctuations in GDP (ie real wages).
• The comovement (eg, consumption and unemployment) can occur with a lag (lags). Some variables (eg productivity and money supply) have a trend that is particularly helpful to predict the next trend GDP (leading indicators). The extent of comovement may be significantly different between countries and over time

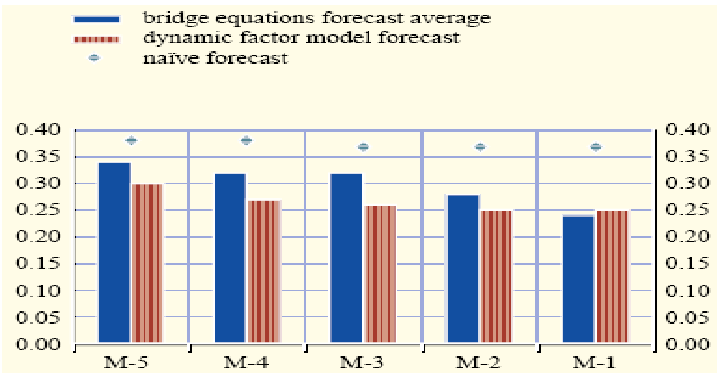
[12] See Quah and Sargent (1993), Forni and Reichlin (1996, 1997, 1998), and Forni and Lippi (1997) .

Finally, factor models can be successfully used for forecasting (Stock & Watson, 1998 and 1999). They deal mainly with predictions in a specification that allows for time-varying factor loadings but not for autoregressive dynamics.

In the article "*The generalized dynamic factor model: identification and estimation*" (Forni et al., 2000), under the dynamic factor model approach[13], provided that a finite number q of eigenvalues of spectral density matrices explode as the dimension of the panel increases (an assumption that seems to be satisfied in most real datasets, with fairly small values of q), the observed series decompose into two unobserved components (a *common component)* and an *idiosyncratic one.* Common components have dimension q, and can be handled by multivariate time series methods, while the idiosyncratic ones, that are only middle inter-correlated, can be treated one by one by univariate methods. Contrary to the static model considered elsewhere in literature, this approach does not impose any restriction on data generating process. They use dynamic factor models to build an index for the European Union. Such index is defined as the common component of real GDP within a model, including several macroeconomic variables for each European country. Factor models can provide a more parsimonious parameterization.

The dynamic factor model-based forecasts tend to be more precise than forecast averages derived from bridge equations at longer horizons. While, for the (final) forecast conducted one month ahead of the GDP data release, the performance of the bridge equations is similar to the one concerning dynamic factor model, as shown in the following chart (*ECB monthly bulletin 04/2008)*. The measure used to value the performance is the RMSE (root mean squared error).

Figure 1.2 Bridge equations versus Dynamic factor model



Source: European Central Bank calculations

One advantage of the dynamic factor model is to obtain a forecast compatible with the model as the simple bridge equations forecasts missing values only on the basis of an univariate autoregressive model.

Some extensions of these approaches have also produced forecasts for demand and value added components as well as for the main euro area economies.

[13] Inspirited by Brillinger's concept of dynamic principal components.

In macroeconomics, factor models are used in business cycle analysis (Forni and Reichlin 1998; Giannone, Reichlin, and Sala 2005, 2006), in the identification of economy-wide and global shocks (Forni et al., 2004-2005; Giannone and Lenza 2004; Giannone and Reichlin 2006), in economic policy (Giannone et al. 2005), in consumer theory (Lewbel 1991), in the construction of indicators and forecasts exploiting information scattered in a huge number of interrelated series (Altissimo et al. 2001), and also in monetary policy applications (Bernanke and Boivin, 2003; Favero et al., 2005).

The DFM, therefore, assumes that GDP growth y_t can be decomposed into a common component χ_t and an idiosyncratic component ε_t; the first captures the bulk of the covariation between growth and a wide range of macroeconomic variables, while the idiosyncratic is assumed to mainly only affect growth:

$$y_t = \mu + \chi_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}(0,\psi) \tag{1.8}$$

where μ is a constant and $\chi_t = \Lambda F_t$, with $F_t = (F_{1t},.....F_{rt})'$ and $\Lambda = (\lambda_1,...\lambda_r)$

The common component is related to growth through a linear combination of a handful of *r* static factors that are, in turn, estimated using information from a large panel of *n* economic indicators, $X_t = (x_{1,t},....x_{n,t})$, where each indicator has a factor representation analogous to that of real GDP growth (see equation 1.8).

It is necessary to consider that in a DFM, the dynamics of static factors can be captured by the following vector autoregressive (VAR) process:

$$F_t = \sum_{i=1}^{p} \beta_i F_{t-i} + Bv_{t,} \qquad v_t \sim \mathrm{N}(0, I_q) \tag{1.9}$$

where the β_i *s* are $r \times r$ matrices, p is the lag length of the process, and B is an $r \times q$ matrix.

In a DFM the number of static factors *r* is assumed to be large relative to the number of common shocks *q,* to capture the dynamics of the economy.

Mariano and Murasawa (2003), Aruoba and others (2009), Boragan and Diebold (2010) suggest extending the dynamic factor model by incorporating data measured at different frequencies. Similarly, Camacho and Perez-Quiros (2010) aim to estimate real GDP growth at a monthly frequency for the euro area by incorporating data on preliminary, advanced, and final GDP releases.

In section 1.3, the application of the DFM will be outlined to analyze smoothed growth by Eurocoin indicator.

1.3 The Dynamic Factor Model and Eurocoin

About eight years ago the Center for Economic Policy Research started publishing and updating monthly Eurocoin, an important application of dynamic factor model. It is an indicator of the Euro area economic activity concerning the medium to long-run growth, published monthly by the Banca d'Italia and CEPR.

New Eurocoin has been recently created (see Altissimo et al. 2006); it is a timely estimate of the medium to long-run component of euro area GDP growth and it has a measure of performance. Since New Eurocoin indicator provides a summary index of the medium to long-run component (MLRG) of the entire GDP for the whole Euro area aggregate, the main contribution of this Thesis is to propose some disaggregated procedures to estimate the following components of European GDP that are developed in chapter 2 and 3:

- Sectoral smoothed growths;
- MLRG concerning some countries belonging to the Euro Area;
- Expenditure components.

In the New Eurocoin (NE) approach, our sectoral and national indicators smooth stationary series by using only contemporaneous values of a large dataset and they are obtained as a linear combination of the smooth factors: the latent factor are "smooth factors", which are generalized principal components of current values of the variables in the dataset. These factors are designed to remove short-run and variable-specific sources of fluctuation.

The Eurocoin construction is based on Generalized Dynamic Factor model, using leading variables in the dataset as proxies for missing future values in the variable of interest. The removal of short-term dynamics of stationary series from the medium to long-run growth can be obtained through band-pass filters that are infinite moving averages and can deteriorate at the end of the sample. Indicators using bandpass filters are based on past and future values of GDP, and they are less reliable in relation to actual data most relevant for economic policy. NE is obtained as a linear combination of a small number of "smooth factors", which are generalized principal components of current values of the variables in the dataset, and they are designed to remove short-run and variable-specific sources of fluctuation. Since only current values of the variables are used, no end-of-sample deterioration occurs.

The *generalized dynamic factor model*, on which *Eurocoin* indicator is based, must have two characteristics: it must be dynamic, because business cycle questions are typically non-static. Secondly, it must allow for cross-correlation among idiosyncratic components, as orthogonality is an unrealistic assumption for most applications. An important feature of this model is that the common component is allowed to have an infinite moving average (MA) representation, so as to accommodate for both autoregressive (AR) and MA responses to common factors. Dynamic factor model is more general than a static-factor model in which lagged factors are introduced as additional static factors, since AR responses are ruled out in such a model. This model

encompasses as a special case the approximate-factor model of Chamberlain (1983) and Chamberlain and Rothschild (1983), that allows for correlated idiosyncratic components but it is static; it generalizes the *exact* factor model of Sargent and Sims (1977) and Geweke (1977), which is dynamic but has orthogonal idiosyncratic components.
The main theoretical tool in this context is Brillinger's theory of *dynamic principal components* (Brillinger, 1981).

In a classic factor model, considering the scalar time series variable Y_t to forecast and let X_t be the N-dimensional time series of candidate predictors, it is assumed that (X_t, Y_{t+h}) admits a factor model with r common latent factors F_t:

$$X_t = \Lambda F_t + \varepsilon_t \tag{1.10}$$

$$Y_{t+h} = \beta_F^{'} F_t + \beta_\omega^{'} \omega_t + \varepsilon_{t+h} \tag{1.11}$$

where ε_t is an $N \times 1$ vector of idiosyncratic disturbances, h is the forecast horizon, ω_t is an $m \times 1$ vector of observed variables (e.g., lags of Y_t) useful, with F_t, to forecast Y_{t+h}. When the idiosyncratic disturbances ε_t are temporarly iid and cross-sectionally independent, we can consider the equation above as a classic model of factor analysis.
Stock, Watson (2002a,b) propose the *static approximate* factor model, in which errors can be correlated either serially or in terms of cross-sectional. They therefore propose a serially correlated version of the approximate factor model of Chamberlain and Rothschild (1983). Under Stock, Watson approach the common component χ_{it} can be expressed as a linear combination $\sum_{j=1}^{r} a_{ij} F_{jt}$ of a small number r of unobserved *static* factors $(F_{1t}, ...F_{rt})$; the loading a_{ij} are real numbers, while all factors are loaded contemporaneously. They prove that in a dynamic factor model the principal components of X_t are consistent estimators of the true latent factors. So, considering the number of time series observations T, the "feasible forecast" Y_{t+h}, based on the estimated factors and coefficients, as N and T $\rightarrow \infty$, converges to the unfeasible forecast that would be obtained knowing factors and coefficients, and then, as N and T grow large, the difference between feasible and unfeasible estimates tends to 0 and Y_{t+h} is asymptotically efficient. The principal factors are then estimated by a non-parametric method, through the principal components.
Eurocoin is obtained as a projection of the bandpassed GDP on factors by a generalization of the principal components concept.

Each series contained in the dataset for the calculation of Eurocoin is obtained as follow

$$x_{it} = \chi_{it} + \xi_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + ... + b_{iq}(L)u_{qt} + \xi_{it} \quad (1.12)$$

with a common component χ and an idiosyncratic ξ, that are orthogonal at all lead and lags.

The common components χ in the classical model of dynamic factors have the following form:

$$\chi_{it} = c_{i1}F_{1t} + c_{i2}F_{2t} + ... + c_{ir}F_{rt} \quad (1.13)$$

Eurocoin is an alternative estimate to c_t, the medium to long run component of GDP.

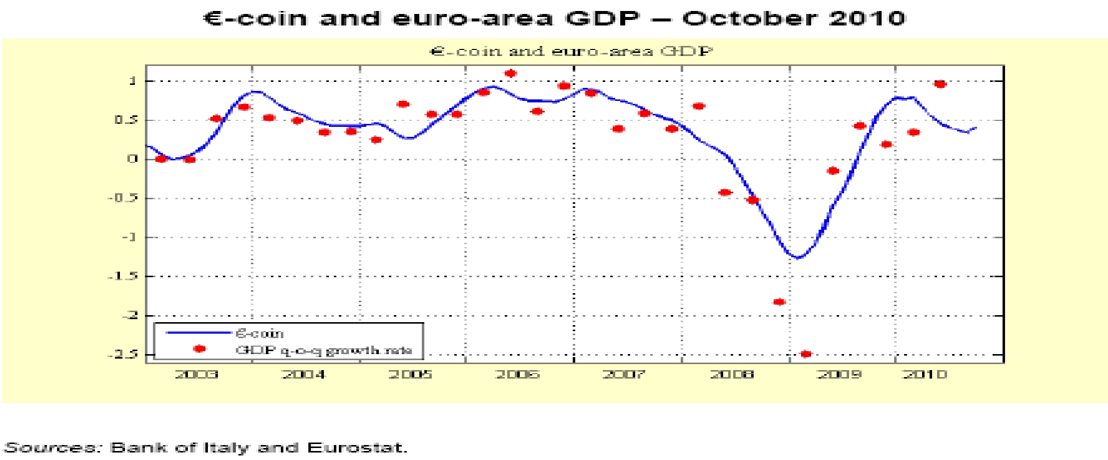
The value of c_t, with the coefficients A_i, at the end of the sample is so estimated[14]:

$$\hat{c}_T = A_1F_{1T} + A_2F_{2T} + ...A_mF_{mT} \quad (1.14)$$

So, New Eurocoin (NE) is the projection of c_t (the bandpassed GDP) on a set of regressors, which are linear combinations of the variables contained in the Thomson Financial Datastream used by the Bank of Italy.

Eurocoin indicator is a monthly published by CEPR and Bank of Italy (Figure 1.3); its estimation is based on the collation of data from an ample set of macroeconomic time series.

Figure 1.3



Source: http://eurocoin.cepr.org/

1.3.1 What Advantages Using Eurocoin ?

The idea that a common-idiosyncratic decomposition provides an important description of variables in business cycle analysis goes back to Burns and Mitchell (1946). The innovation of Eurocoin with respect to the econometric literature is a procedure to remove both the

[14] See *New Eurocoin . A tutorial Note*; http://eurocoin.bancaditalia.it4

idiosyncratic ad the short-run components, so that the resulting factors are both common and smooth.

The estimate of underlying monthly growth trends by **€-coin approach** has four key advantages:

(i) monthly frequency and timely estimate, released several months ahead of the official euro-area GDP estimate;

(ii) immunity to most erratic components and net of the more volatile components (seasonal variations, measurement errors and short-run volatility, local and sectoral shocks). Similarly, by the disaggregated analysis that we carry out in chapter 2 and 3, obtained by projecting national and sectoral components of growth on common factors, we eliminate local and sectoral shocks, to estimate in real time:

- National smoothed GDP;
- Sectoral smoothed growth rate;
- to determine the contribution of different medium to long-run components (consumption, investment, exports, imports) of GDP in the European business cycle

(iii) Eurocoin smoothes stationary series by using only contemporaneous (current) values of a large dataset so that no end-of-sample deterioration occurs.

(iiii) Eurocoin can't be defined as an estimate of a latent variable, being different in this respect from the coincident indicators constructed e.g. in Stock and Watson (1989) or the one produced by OECD. It is a real time estimate of the medium to long run component of GDP growth, and the latter is observable, although with a long delay. Therefore, the performance of this indicator and the ones concerning the disaggregated models that we build in this thesis (by implementing Eurocoin approach) can be measured.

More precisely, the value of the target (the bandpassed data) is available with good accuracy only at time T +h, for a suitable h. Therefore, our indicators produced at time T can be compared with the target at T produced at time T + h.

In fact, isolating the business cycle in integrated series is a well known result in literature, so that bandpass filters can therefore deteriorate at the end of the sample, when the missing values useful to describe the cycle is replaced by predictions. Bandpass filters can eliminate erratic components, and they are based on past and future values of GDP. However, contrary to the Eurocoin, they are less reliable for most recent data, very relevant for economic policy. Below is a diagram that shows what happens at the end of the sample, both projecting European GDP on French factors (using factors only built by French variables as we will explain more deeply in chapter 2 by Eurocoin methodology (generalized dynamic factor model – figure 1.4) and filtering French GDP by Baxter and King or moving average filter (see figures 1.4/A and 1.4/B,C). We name FRANCOIN the filtered data that we outline using French factors (in figures 1.4/A), in green.

Figure 1.4/A EUROCOIN versus FRANCOIN

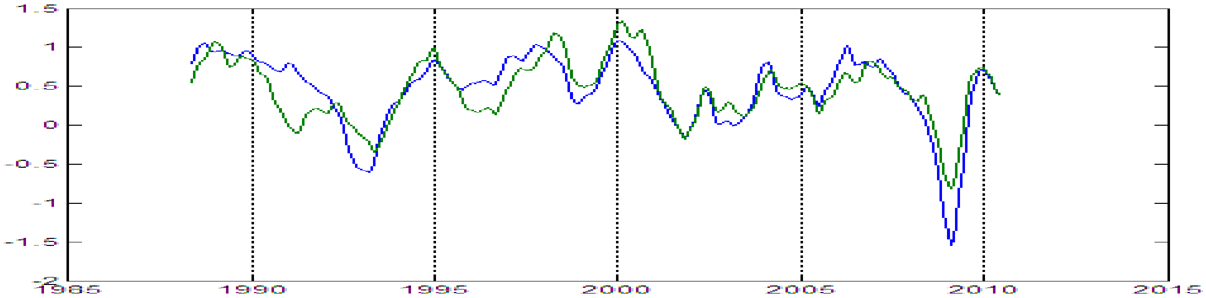


Figure 1.4/B : Filtering French GDP by Baxter and King

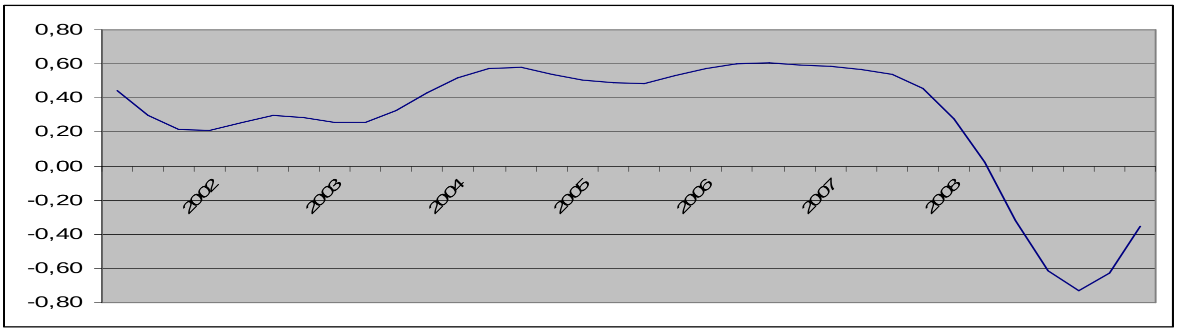
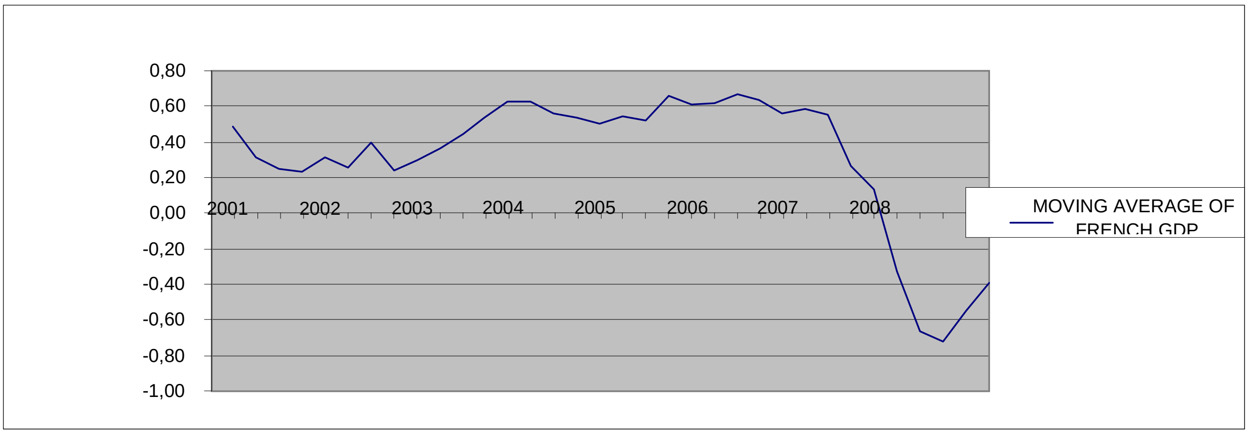


Figure 1.4/C : Filtering French GDP by a centered MA



Using FRANCOIN indicator built by the Eurocoin approach and based on generalized dynamic factor model (GDFM), it is possible to analyze the absence of deterioration at the end of the sample as it uses more recent data information concerning GDP.

On the contrary, both moving averages and Baxter-King filters are not reliable at the end of the sample. However, BK is less erratic than moving average.

1.3.2 Transforming the Common Factors in the Eurocoin Approach

An important step in building a real time indicator based on the GDFM is to have some consistent estimates of the common components.

Under equation (1.11), different consistent estimators have been proposed for the space G_F spanned by the factors F_{jt}, see Stock and Watson[15] (2002a, 2002b), and Forni, Hallin, Lippi, Reichlin (2005), as both the number of observations in each series and the number of series in the dataset tend towards infinity. Consistent estimates of the common components χ_{it} are obtained by projecting the variables x_{it} on the estimated factors.

Supposing the variables in the dataset are obtained as first differences, $x_{it} = X_{it-1} - X_{it-1}$, so that the factors $F_{1t}, F_{2t}, \ldots, F_{rt}$ are linear combinations of first differences.

Assumption proposed in the generalized dynamic factor model, building the Eurocoin, is that the variables x_t, as well as the GDP are driven by the factors F_{kt}. The month-on-month change of the GDP (unobservable) can be outlined as follows:

$$x_t = (1-\mathrm{L})\, X_t = \chi_t + \xi_t = \chi_t = c_1 F_{1t} + c_2 F_{2t} + \ldots.. + c_r F_{rt} + \xi_t \qquad (1.15)$$

with ξ_t orthogonal to the factors.

On the other hand, y_t is a quarter-on-quarter rate of change, whereas the variables $x's$ used to construct the factors are month-on-month rates of change, so that representing y_t in terms of the factors transformed by

$(1 + \mathrm{L} + \mathrm{L}^2)^2$, i.e.

$$y_t = (X_t + X_{t-1} + X_{t-2}) - (X_{t-3} + X_{t-4} + X_{t-5}) = (1 + \mathrm{L} + \mathrm{L}^2)^2\, x_t = c_1[(1 + \mathrm{L} + \mathrm{L}^2)^2 F_{1t}] + c_2[(1 + \mathrm{L} + \mathrm{L}^2)^2 F_{2t}] + \cdot\cdot\cdot + c_r[(1 + \mathrm{L} + \mathrm{L}^2)^2 F_{rt}] + (1 + \mathrm{L} + \mathrm{L}^2)^2\, \xi_t \qquad (1.16)$$

Then, assuming that our Thomson Financial Datastream can be modelled as a dynamic factor model, the projection of c_t (e.g. the bandpassed sectoral or national growth rate that we use in this thesis) on factor space will be estimated by using transformed regressors:

$$(1 + \mathrm{L} + \mathrm{L}^2)^2 F_{kt} \qquad\qquad k = 1, \ldots, r;$$

the same transformation will be applied to smooth factors.

NE performs equally well within and at the end of the sample. It outperforms the truncated band-pass filter at the end of the sample, both in terms of turning-point signaling and fitting.

[15] Stock and Watson use the first r principal components of the variables x_{it}.

1.3.3 The Number of Smooth Factors in Eurocoin Methodology

The second main step in building our real time indicators to track the underlying sectoral and national growth rates (see chapter 2 and 3), is choosing of the number of factors useful to the smoothing of the 157 time series (belonging to the main European countries) contained in the dataset TFD.

Let x_t, χ_t and ξ_t be the vectors of the variables x_{it}, their common and idiosyncratic components respectively; ϕ_{it} is the medium to long-run component of χ_{it}, then:

$$\phi_{it} = \beta(L)\chi_{it} \tag{1.17}$$

$$\psi_{it} = \chi_{it} - \phi_{it} \tag{1.18}$$

For the spectral density matrices:

$$S_x(\theta) = S_\chi(\theta) + S_\xi(\theta) = S_\phi(\theta) + S_\psi(\theta) + S_\xi(\theta) \tag{1.19}$$

For estimates of S_χ and S_ξ see Forni, Hallin, Lippi , Reichlin (2000).

Integrating over the interval $\begin{bmatrix} -\pi & \pi \end{bmatrix}$ it follows this decompositions of the variance-covariance matrix of the $x's$:

$$\sum_x = \sum_\chi + \sum_\xi = \sum_\phi + \sum_\psi + \sum_\xi \tag{1.20}$$

Regarding the construction of New Eurocoin, we observe that the number of smooth factors can be determined by the following procedure:

1) Firstly, we estimate q, the dimension of the white noise u_t in each series x_t of the dataset (see 1.10), and it is computed by the criterion proposed in Hallin, Liska (2007). They develop information criterion for determining the number *q* of common shocks in the *general* dynamic factor model, based on the fact that this number *q* is also the number of diverging eigenvalues of the spectral density matrix of observations as the number *n* of series goes toward infinity. They provide sufficient conditions for the consistency of the criterion for large *n* and *T* (where *T* is the series length), showing by simulating its good finite-sample performance. Hallin and Liska attempt to fill a gap on dynamic factor models in literature by providing an efficient and flexible tool for identifying the number, *q*, of factors. Application to real data suggest that the number of factors driving the U.S. economy may be lower than suggested by an application of static or restricted dynamic factor models.

2) Secondly, based on the determination of q = 2, to determine smooth factors, by (1.17) we estimate *covariance matrices* $\sum_{\chi} \sum_{\phi} \sum_{\xi}$ concerning common components and its medium to long-run component and the idiosyncratic ones respectively;

- $\hat{\sum}_{\chi}$ and $\hat{\sum}_{\xi}$ are obtained by integrating the spectral density $\hat{S}_{\chi}(\theta)$ and $\hat{S}_{\xi}(\theta)$ over the interval $\left[-\pi \quad \pi\right]$;
- $\hat{\sum}_{\phi}$ is computed by integrating $\hat{S}_{\chi}(\theta)$ over $\left[-\pi/6 \quad \pi/6\right]$.

3) Thirdly, the dimension *r* of the factor space $G_{F,}$ is estimated using the criterion proposed by Bai and Ng (2002) as method for identifying it. This criterion is based on information-theoretical quantity. They establish under appropriate assumptions, the consistency of their static method as *n*, the cross-sectional dimension, and *T*, the length of the observed series, both tend towards infinity. This criterion has been adapted to the restricted dynamic framework (Bai and Ng 2005; Amengual and Watson 2006), but, in the general dynamic case, this approach can overestimate the *r*.

Using Bai and Ng's method for the construction of Eurocoin, we determine that the dimension of the factor space *r* = 12. Projecting c_t, the medium to long run component of GDP, on the first *r principal components* of the series in the dataset, which is a basis of the factor space, we obtain the projection κ_t. Principal components are filtered with $(1+L+L^2)^2$, see (1.16).

4) The generalized eigenvectors $\upsilon_k, k=1,...,r$ associated with the generalized eigenvalues $\lambda_1,....,\lambda_r$, are ordered from the largest to the smallest; they satisfy the following conditions:

$$\hat{\Sigma}_{\phi}\, \upsilon_k = \lambda_k (\hat{\sum}_{\chi} + \hat{\sum}_{\xi})\upsilon_k \tag{1.21}$$

with the normalization constraints

$$\upsilon_k^{'}(\hat{\sum}_{\chi} + \hat{\sum}_{\xi})\upsilon_k = 1 \tag{1.22}$$

see Anderson (1984, Theorem A.2.4, p.590)

The associated *generalized principal components* (GPC) are so calculated as follows:

$$W_{kt} = \upsilon_k^{'} x_t \tag{1.23}$$

Projecting c_t, the medium to long run components of GDP, on the first s generalized principal components of the series in the dataset, we obtain the projection $\kappa_t^{(S)}$.

Generalized principal components (GPC) are filtered as in (1.14).

Lastly, let ρ and ρ_s be the R^2 of the regressions obtained by the projection of c_t on κ_t and $\kappa_t^{(S)}$ respectively: starting with $s=1$, we stop when the difference between the two R^2 is approximately of the same size: it is called $\bar{s}$ the number of GPC (smooth factors) so computed. It is possible to verify that $\bar{s}$ is smaller than r and that the projection of c_t on the $\bar{s}$ smooth factors (that is Eurocoin indicator) is smoother than κ_t.

1.3.4 Determining the Number of Dynamic Factors: Alternative Criteria in Literature

Alternative criteria based on the theory of random matrices have been developed by Kapetanios (2005) and Onatski (2006), for the number of static factors but in a model with iid idiosyncratics (instantaneous idiosyncratic cross-sectional correlation is allowed in Onatski where, however, Gaussian assumptions are made).

Bai and Ng (2007) suggest a two-step procedure for determining the number of dynamic factors in factor models.

The procedure in Bai and Ng (2007) relies on the fact that the $r \times r$ matrix of innovations to the static factors (see equation 1.7.2) has rank equal to the number of dynamic factors *q* (the common shocks).

The first step of the procedure requires that the number of static factors *r* is determined by the information criteria described in Bai and Ng (2002). Then, once the number of static factors *r* is set, the rank of spectrum of *q* dynamic factors is estimated by the eigenvalues of residual covariance (or correlation matrix) of VAR in the *r* static factors.

Matheson (2011) explains that Bai and Ng criteria (2002) generally produced too many factors, deteriorating the forecasting performance of the DFM. Likewise, he explains that the approach used by Giannone et al. (2005) and Matheson (2010), where the number of factors explains a certain percentage of the variation in a few key series, was not well suited to a multi-country setting, because there is significant variation in the explanatory power of the factor model across these countries.

However, following Stock and Watson (2002), the number of factors can be determined minimizing Schwarz's Bayesian information criterion (SBC); in Matheson (2011) the number of common factors *r* is chosen by regressing quarterly real GDP growth on the common factors for *r* = 1*; :::;* 8; the number of factors is the one that minimizes the SBC.

1.4 Among Bridge Equations and Dynamic factor models

The EuroCoin indicator is one of the first attempts to publish an economic indicator using a large panel of data in real time. Since the variables in the dataset (see annex 3) for a description of the Thomson Financial Datastream used in this Thesis are often available with different delays, in order to solve this end-of-sample unbalance problem, Eurocoin indicator uses the following approaches:

- shifting forward the variables that are available with delay;
- the variables that are not available might be predicted using ARMA models or the EM algorithm.

An alternative way to exploit the information to fill the large monthly and quarterly variables (which will not be used in this research), was proposed by Giannone et al. (2005), first applied on data from the United States at the Board of Governors of the Federal Reserve and now regularly carried out at the ECB (European Central Bank). This method aims to combine predictors in few common factors that are then used as regressors in bridge equation by the Kalman filter (Bridging with factors - *BF method*).

In Giannone et al. (2008), it is possible to find advantages of BF model that outperforms the AR(1) and the BES model (of Diron), exploiting richer information set of cross correlations by multivariate information and the more extensive use of survey and financial market data; for BF model there is a clear decline of RMSE with increase in monthly information.

While the EuroCoin indicator uses some approaches to handle missing data at the end of samples that change with the pattern of available data, the approach proposed by Giannone et al. (2008) and followed, for example, by Barhoumi et al. (2008) and Matheson (2010), uses the Kalman filter to make estimates in the presence of missing observations at the end of the sample due to publication lags, the so-called “jagged edge”[16].

It is important in this Thesis, therefore, to describe this method as an implementation of bridge equations by dynamic factor model and, similarly, to generalized dynamic model in Eurocoin, it relates monthly series to a vector of latent (common) factors plus an idiosyncratic component. Nevertheless, we will not use Kalman filter for our goal. In implementing Eurocoin methodology in chapter 2 and 3, it is however useful to apply spectral filters as underlined in *New Eurocoin, A tutorial Note* (Bank of Italy)[17].

Consider the vector of n stationary monthly series $x_t = (x_{1,t},...., x_{n,t})'$, $t = 1,...,T$, which have been standardized to have mean zero and variance one. The *dynamic factor model* considered by Giannone et al. (2005), above named BF model, is then given by the equations

$$x_t = \Lambda f_t + \xi_t \qquad\qquad \xi_t \sim \mathrm{N}\left(0, \Sigma_\xi\right)$$

[16] However, use of the Kalman filter is limited in short samples (see Stock and Watson, 1988) and some a priori judgments should be made concerning the data gene rating process.
[17] http://eurocoin.bancaditalia.it

$$f_t = \sum_{i=1}^{p} A_i f_{t-1} + \varsigma_t$$

$$\varsigma_t = \beta\eta_t \qquad \eta_t \sim \mathrm{N}(0, I_q)$$

From an $n \times r$ matrix of factor loadings, equations described above relates monthly series x_t to a $r \times 1$ vector of latent factors $f_t = (f_{1,t,...} f_{r,t})'$ plus an idiosyncratic component $\xi_t = (\xi_{1,t}, ..., \xi_{n,t})'$. The latter is assumed to be multivariate white noise with a diagonal covariance matrix $\sum_\xi$. Equation concerning f_t describes the law of motion for the latent factors f_t, which are driven by q-dimensional standardized white noise η_t, where B is a $r \times q$ matrix, with $q \leq r$; $\varsigma_t \sim N(0, BB')$. $A_{1,}....A_p$ are $r \times r$ matrices of parameters, and it is assumed that the stochastic process for f_t is stationary. Defining quarterly GDP growth as the average of monthly latent observations: $y_t^Q = (y_t + y_{t-1} + y_{t-2})$, it is possible to obtain quarterly factors as $f_t^Q = (f_t + f_{t-1} + f_{t-2})$.

The factors-based bridge equation is:

$$\hat{y}_t^Q = \beta' \hat{f}_t^Q$$

where β is an $r \times 1$ vector of parameters. In the 3rd month of each quarter, the forecast for quarterly GDP growth, $\hat{y}_t^Q$, is evaluated as the quarterly average of monthly series:

$$\hat{y}_t^Q = \frac{1}{3}(\hat{y}_t + \hat{y}_{t-1} + \hat{y}_{t-2})$$

and the forecast error is thus evaluated $\varepsilon_t^Q = y_t^Q - \hat{y}_t^Q$. The errors $\varepsilon_t^Q \sim N(0, \sigma_\varepsilon^2)$ and they are mutually independent at all leads and lags. To insure consistency among monthly indicators and quarterly GDP, all monthly variables are transformed to insure that the corresponding quarterly amounts are defined by $x_{it}^Q \sim (x_{it} + x_{it-1} + x_{it-2})$ where $t = 3k$ and $k = 1.....,T/3$: so, series in differences enter the factor model in terms of 3 month changes, having defined quarterly GDP growth as the three month average of monthly latent observations. Monthly panel is not balanced at the end of the sample due to different publication lags of the data. The fact that dynamic factor model deals easily with the unbalanced data is an important property of it.

Kalman filter (1960) was used at the beginning for the analysis of signals and to make inferences about unobservable components. It has been used in econometric applications since 1980. The statistical and econometric literature, which is interested in the estimation of unknown parameters, shows how to apply Kalman filter to calculate the likelihood function. This tool is a system of equations that allows to update estimators. *State space* is a very general form of representing dynamical systems, and to make inferences about unobservable components. Kalman filter is a recursive estimator for making inference about unobservable components likewise a recursive estimator. It has two distinct phases: Predict and Update. The first uses state estimates from previous time step to produce an estimate of state at the current time step. This predicted state estimate is also known as the *a priori* state estimate as, although it is an estimate of the state at the current time step, it does not include observation information from the current time step. In the update phase, the current *a priori* prediction is combined with current observation information to refine the state estimate. This improved estimate is the *a posteriori* state .

From the state space form of the model, Kalman filter techniques can be easily applied to obtain efficient forecasts of GDP growth y_Q^t from the unbalanced data sets. It is applied to the state space representation of this model. The purpose of the Kalman filter is to recursively construct the estimate, e.g. concerning state β_{t+1}, denoted by $\hat{\beta}_{t+1}$, given the observation R_{t+1} by updating the estimate of the state obtained immediately before and denoted by $\hat{\beta}_t$ (see Harvey, 1981).

1.5 Approaches to Business Cycle Measurement

NE provides an index of the current economic situation with the purpose of tracking the underlying GDP (the MLRG) for the entire Euro area, thus it is an "*aggregate indicator*".

The main aim of this thesis is to propose a new theoretical framework for disaggregated business cycle analysis *by sectors, countries and expenditure components,* based on the Eurocoin approach. In literature concerning the Eurocoin indicator, real time and ex-post performance are generally studied with respect to a measure of the "trend-cycle GDP growth" obtained in the middle of the sample by a band pass bilateral filter on GDP growth (see section 3.2.1).

Unlike from the definition of MLRG, in a cycle, even the oscillations of a period longer than 8 years are usually removed (for different definitions of the cycle, see Stock and Watson, 1999).

With regard to the main definitions of "business cycle", it is however useful to distinguish between a classic and a modern approach.

The *classic approach* embodied in the original NBER methodology were based on the concept which focuses on fluctuations in the absolute level of economic activity.

This approach[18] is based on business cycle analysis following the *NBER* (National Bureau of Economic Research) Business Cycle Dating Committee that identifies phases of recovery and recession in relation to variables that are expressed in levels, including long-term trend. It dates back to the definition originally proposed by Burns and Mitchell (1946)

In the NBER approach, which is still utilized in the U.S. to establish business cycle official chronology, the dating of reference cycle is based on the classical definition. Upturns and downturns are identified with respect to absolute levels of variables used in the analysis, without removing their secular component.

The assumption of this method is that the system of economy has cyclical trends of expansion and recession, when most parts of its activities are in the same situation.

In synthesis, following *NBER methodology:*

- Business cycles are fluctuations in an aggregate economic activity, persistent over time and widespread across sectors. They are meant to be «recurrent but not periodic», with a duration ranging from «more than one year to ten or twelve years».
- The cycles are defined as "common movements" across economic processes identified by downturns and upturns in many economic variables «concentrated around certain points in time»
- recessions are defined as periods of low (or even negative) growth, and expansions as periods of high growth.

[18] This classic definition is followed by OECD in specifying growth cycle outlook by CLI (Composite Leading Indicators),

Concerning the *modern approach*es, they are often based on the «growth cycle» definition. The cycle is identified with the deviation of economic activities from their long-term trend. With regard to these new approaches to business cycle measurement, we highlight the comovement tenet that was endorsed by Lucas (1977) and, on empirical grounds, by Sargent (1987) and Stock and Watson (1989).

Sargent (1987) provided a reinterpretation of the Burns and Mitchell approach based on frequency domain concepts: «the following definition seems to capture what experts refer to as the business cycle. Business cycle is the phenomenon of a number of important economic aggregates (..) being characterized by high pairwise coherences at low business cycle frequencies».

In Stock and Watson (1989) the co-movement definition[19] was made operative through a *dynamic latent factor model* similar to the one we will use in our research, where we focus on medium to long-run components of growth (MLRG), i.e. the smoothed components of GDP growth rate obtained by removing the fluctuations of a period shorter than or equal to one year. MLRG bears no relationship to any definition of trend. In New Eurocoin, the *latent factor* are "smooth factors". They are generalized principal components of current values of variables in the dataset

Stock and Watson (1990 and 1998) produced an interesting body of empirical work that, in line with the original NBER research project, focuses on the description of the business cycle properties of a large set of indicators.

A reference technique for the detection of *turning points* is represented by the procedure of *Bry and Boschan* (1971*)*. It is a nonparametric algorithm that can be applied to a single monthly time series[20], adjusted for seasonality. It consists in the extraction of the points identified as local maxima/minima and satisfying some constraints. Each complete cycle should last at least fifteen months, and every single phase of the cycle (with alternation of expansions and recessions) should last, at least, five months. Turning points too close to the extremes of ranges are discarded. This procedure will be discussed in detail in Annex 1, and it will be used for empirical applications in chapter 3.

In the U.S. tradition, the identification of turning points is also judgmental and based on the thorough analysis of a small but comprehensive set of indicators (including output, employment, sales and income variables) representing the evolution of economic activities (Moore and Zarnowitz, 1986, and Zarnowitz, 1992).

CEPR (Centre for Economic Policy Research) definition of recession, expansion, peaks and troughs correspond to the so called "growth cycle". Recession, for example, is defined as a prolonged period of declining growth in the cyclical component of GDP (as measured by the movements of EuroCOIN). Analogously, an expansion is a prolonged period of increasing

[19] See also Diebold and Rudebusch (1996).

[20] An extension to quarterly series is proposed in Harding and Pagan (2002).

growth. Troughs and peaks are defined as the ending points of expansions and recessions respectively, i.e. as points of minimal or maximal growth. Expansions as defined from CEPR anticipate classical expansions. According to this definition, the euro area economy can be in an expansionary phase even in periods when GDP is declining, provided that growth is increasing.

In this thesis, recession (expansion) is defined by applying Bry and Boschan procedure to the medium to long run components of growth rate as measured by the movements of EUROCOIN and of disaggregated indicators mentioned in chapters 2 and 3 in terms of quarterly changes. Expressing indicators in terms of quarterly changes could be useful to determine a cyclical turning point.

In section 3.2 we define a *turning point* as an upturn or a downturn, then as a slope sign in the target that we use to compare our disaggregated indicators.

Burns and Mitchell (1946) affirmed that "a cycle consists of expansions occurring at about the same time in many economic activities". So one of the characteristics of the cycle is represented by the co-movements among variables. We could have an extension of the Bry-Boschan procedure to a multivariate framework. A possible solution relies on an *indirect* approach, in which the turning points detected on a number of single series are aggregated following some specified rules. To this aim we show disaggregated indicators concerning economic activities by sectors, countries and expenditure components of GDP.

1.5.1 Relationships among Business Cycles and Variables

The analysis that we carry out in next chapters will also determine the contribution of the medium to long run expenditure and national components of GDP to the European business. A point of reference for this type of analysis is the work of Stock, Watson (1998) [21] , that examined the empirical relationships in postwar USA between the aggregate cycle and cyclical components of individual time series (Consumptions, Investments, Stocks, Foreign Trade), and whether individual series are useful in predicting aggregate fluctuations.

Giannone, Lenza, Reichlin (2009) analyzed the characteristics of euro area business cycle over the last forty years, typically performed with quarterly data. They made the choice of using annual data as quarterly data are not available for all countries for a sufficiently long sample (they have been harmonized only since 1991). First, they look at national cycles, and then their dynamic relation over time.

Below are the main descriptive statistics they use to describe national business cycles:

[21] J. Stock and M. Watson (1998).

- percentage difference between the log of real GDP per head $Y_{i,t}$ of each country and the euro area aggregate in different years and sub-periods:

$y_{i,t} = y_{ea,t} + (y_{i,t} - y_{ea,t})$ where $y_{ea,t}$ refers to the euro area. (1.24)

Since each country's growth depends on both euro area developments and its idiosyncratic dynamics, they consider the following decomposition:

$\Delta y_{i,t} = \Delta y_{ea,t} + (\Delta y_{i,t} - \Delta y_{ea,t})$ y$\Delta y_{i,t} = \Delta y_{ea,t} + (\Delta y_{i,t} - \Delta y_{ea,t})$ (1.25)

where Δ is the difference operator.

The variations in the gap $(\Delta y_{i,t} - \Delta y_{ea,t})$, that is the growth differential with respect to the euro area, represent country specific business cycle developments that can originate either in idiosyncratic shocks or in heterogeneous reactions to shocks concerning euro area. It is a simple measure of business cycle heterogeneity.

The cycle amplitude, for different variables, is typically measured by standard deviation, estimated by the standard error which is a measure of variability of the estimator.

For instance, the relative contribution of household consumption to GDP cycle, in literature, is generally[22] calculated as the ratio concerning covariance between cyclical components of GDP and consumption with GDP cycle variance:

$$contributo(cons) = \frac{Co\,\mathrm{var}(pib\_c, conso\_c)}{Var(pib\_c)}. \quad (1.26)$$

Results concerning the application of 1.26 to the smoothed components of the Euro Area GDP are presented in section 2.5.

[22] See Fournier, J.Y. (1999)

1.6 The Combination of Forecasts

In chapter 2, section 2.4, we show a combination (using regression method to determine the relative weights) of real and financial MLRG, useful to analyze the impact of real and financial data in estimating smoothed GDP. We will divide the 157 variables contained in Thomson Financial Datastream (TFD), that compose our Real Time Dataset, in real and financial variables. Main econometric techniques used for this type of experiment are presented in this section. Generally, a single forecasting model cannot be fully appropriate for all situations. Even if it proved to be satisfactory in terms of Root Mean Squared Forecast Error (RMSFE) (or some other criteria) for a specific situation, its performance could change according to different sample realizations. In this case, an improvement of predictive ability can be achieved by forecasting combination. The idea of combining forecasts from different models is not novel in literature. A historical review on information combination can be found in Bates and Granger (1969), Clemen (1989), Yang and Zou (2004), Yang (2004) and Timmermann (2005). Moreover, as remarked by Kennedy (2001), there is some agreement in literature that the 'best forecasting method' is a `combined forecast', formed as a weighted average of a variety of forecasts, each generated by a different technique. In this research we will use a forecasting combination defining appropriate weights for two different nowcasts of medium to long run GDP growth (obtained respectively by real and financial variables) (see chapter 2). There are alternative methods to compute the weights useful for combining several forecasts in a single forecast:

a) a first is the *simple average scheme (EW)*, assigning equal weights to each individual forecast. This method represents a benchmark against which more refined combining solutions can be evaluated, even if it is often the best solution in empirical exercise.

b) A second is regression, that we will use in this research;

c) another method is algorithm called *AFTER (Aggregated Forecast Through Exponential Reweighting)*, proposed by Yang (2001). The weights are determined on the basis of the history of errors made by the individual models generating the forecasts.

We consider the case of J individual forecasts $\left(\hat{y}_{t,1}, \hat{y}_{t,2}, \ldots \hat{y}_{t,j}\right)$; let $W_{1,j} = 1/J$, implying equal weights in the first period. For $t \geq 2$ the weights are given by

$$W_{t,j} = \frac{\prod_{i=1}^{t-1} \hat{\sigma}_{j,i}^{-1/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{t-1} \frac{(Y_i - \hat{y}_{i,j})^2}{\hat{\sigma}_{j,i}}\right)}{\sum_{j'\geq 1} \prod_{i=1}^{t-1} \hat{\sigma}_{j,i}^{-1/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{t-1} \frac{(Y_i - \hat{y}_{i,j})^2}{\hat{\sigma}_{j',i}}\right)} \tag{1.27}$$

where $\hat{\sigma}_{j,i}$ is the estimated error variance of model j based on information up to i. The weights are constrained to lie between range $[0,1]$; the more a weight approaches 1 the larger the ability of the corresponding model to forecast the actual value in all previous periods. It is also possible to consider *non linear combinations* or approaches based on the *variance-covariance* matrix.

1.6.1 Combining GDP Estimations: main Properties

Economic forecasts are often based on a limited set of information, therefore, in this thesis, we use Thomson Financial Datastream that contains 157 real and financial variables concerning different European countries. Assuming a large sample and linearity, a single forecast could be considered as a good approximation of the optimum. However, any economic forecast cannot be the best that could possibly be achieved given all the information in the universe (Granger and Newbold, 1977).

So, a way to improve forecasting accuracy is to combine more forecasts of the same quantity. We will obtain a combination of forecasts that, as we show in this section, cannot do worse than the best individual nowcast.

Considering the simple combination of two individual forecasts, let $y_{t,1}$ and $y_{t,2}$ be forecasts of the same variable Y_t with forecast errors

$e_{t,j} = Y_t - \hat{y}_{t,j}$ j=1,2,…..

such that $E(e_{t,j}) = 0$

$$E(e_{t,j}^2) = \sigma_j^2$$

$$E(e_{t,1}e_{t,2}) = \rho\sigma_1\sigma_2$$

If we consider the following combination by a weighted average:

$$\hat{y}_{t,c} = k\,\hat{y}_{t,1} + (1-k)\,\hat{y}_{t,2} \tag{1.28}$$

the error will be equal to:

$$e_{t,c} = Y_t - \hat{y}_{t,c} = ke_{t,1} + (1-k)e_{t,2} \tag{1.29}$$

and the variance error will be

$$\sigma_c^2 = k^2\sigma_1^2 + (1-k)^2\sigma_2^2 + 2k(1-k)\rho\sigma_1\sigma_2 \tag{1.30}$$

It is possible to show that the minimum of the variance error (1.25) is less than the error variances of individual forecasts, unless either ρ is exactly equal to σ_2 / σ_1 or to σ_1 / σ_2. As $\rho, \sigma_1^2, \sigma_2^2$ must be estimated. Being unknown, optimal combination can never be achieved.

Newbold and Granger (1974), Winkler and Makridakis (1983) show that ignoring correlations between the forecasts can improve the accuracy of the combined forecasts (particularly for small samples).

1.7 Aggregation with Time-Varying and Stochastic Weights

As we have described in chapter 1, New Eurocoin (NE) is the projection of c_t (the whole Euro bandpassed Gross Domestic Product) on a set of regressors, which are linear combinations of variables contained in the *TFD* used by the Bank of Italy to estimate the underlying monthly growth trends. Since the New Eurocoin indicator provides a summary index of medium to long-run component (MLRG) of GDP only for the entire Euro area aggregate, it is interesting to know more about the theoretical properties of forecasts based on aggregate or disaggregate information (e.g. growth rate by country). In this section, we investigate the properties of forecasts for aggregates both with stochastic weighting schemes and fixed base period exchange rates.

Official data for European area considered have existed only since 1999 while, for the pre-1999, we have only national series available. To perform an analysis, it is therefore necessary to first aggregate national data in order to construct series for the euro area, and then merge the series obtained for Eurostat. It is fair to say that there is no indisputable aggregation method for data pre-1999. Furthermore, it is necessary to understand what the most appropriate methodology is in order to link these observations up with the most recent ones provided by Eurostat. Existing literature has proposed several methods of aggregation. Thus, in Monticelli and Strauss-Kahn (1992) data are processed in a single currency using the bilateral exchange rate with the ECU and then aggregated. While in Fase and Winder (1998), they are processed using the fixed exchange rates. Fagan and Henry (1999), however, propose to calculate the aggregate data directly as a geometric weighted averages of national data, without converting them into a single currency. Thus, for each variable x (in logarithm), the aggregated data will be obtained as $X_z=\sum_z w_z X_z$ following a scheme of fixed weights w based on GDP at purchasing power parity (PPP). The problem is that using current rates causes a "distortion" in the dynamics of the series. Growth rate, standard deviation, and all the moments in general are influenced by movements in exchange rates by introducing a large number of spurious shocks in series. It is also true that by analyzing structural relationships between macroeconomic variables for fixed exchange rates is to assume that it is not a relevant variable, which is true for the post-1999, but not for the periods pre - 1999.

Monticelli and Tristani (1999) use a system of weights given by the relative magnitude of the output in 1993 (rates PPP exchange) and Coenen and Vega (2001), Coenen and Wieland (2000), Gerlach and Svensson (2003) and Peersman and Smets (2003) use data taken from the *Area-wide model* (AWM) database Fagan et al. in which weights used in the aggregation of individual series is income at constant market price (PPP) in 1995.

For the Euro Area GDP, in chapter 2 we also use data from Fagan et al. (2001) until the first quarter 1991, when we build the main figures concerning ex post indicators by country and expenditure components. From then on, we use official Eurostat and ECB series (rescaling data prior to 1992 to avoid a sudden change in level).

And the problem remains of fitting this part of the series with the data obtained by Eurostat. The criterion by which tying the series is directly suggested by Fagan, Henry and Mestre and consists in making a backdate of the "Eurostat" series using the growth rates contained in the AWM database.

As the weights are strictly related to the working populations in the different member states, it may be more appropriate to view the weights as stochastic. There are many series where aggregation is constructed by stochastic weights. For example, there are several proposals for aggregating GDP growth rates based on stochastic weights.

Suppose y_t^{EU} *t* denotes euro-area GDP and $y_t^{(i)}$ is the corresponding figure for country *i*. A good review of possible aggregation methods used in literature focusing on aggregating the level series is given in Winder (1997). Winder is in favour of using fixed base period exchange rates because as it avoids the problem of introducing an unwanted component, the (de-) appreciation of exchange rate in the growth rate series.

Then Winder (1997) computes the EMU growth rate as

$$\Delta \log y_t^{EU} = \sum_{i=1}^{n} \frac{y_{t-1}^{(i)} \Big/ e_{TB}^{(i)}}{y_{t.1}^{EU}} \Delta \log y_t^{(i)} \tag{1.31}$$

where $e_t^{(i)}$ denotes the exchange rate of country *i* in period *t* and *TB* signifies a fixed base year. For Winder the weights are:

$$\omega_{it} = \frac{y_{t-1}^{(i)} / e_{TB}^{(i)}}{y_{t-1}^{EU}}$$

Exchange rate is necessary when an aggregate series for pre-EMU period is constructed. Given the substantial relative price-changes due to currency de- and revaluations, using for instance 1979 as base year would lead to aggregate series that are different from the aggregate series constructed using 1999 as base year.

A useful property of the constructed aggregate level series, $x_t^{EU} = \sum_i x_t^i / e_T^i$, is that the corresponding growth rate series is a weighted average of each country’s own growth rate, with weights being the respective current shares of the countries in the aggregate series, i.e.

$$\Delta \ln(x_t^{EU}) \cong \frac{\Delta x_t^{EU}}{x_{t-1}^{EU}} = \frac{\sum_i x_t^i / e_T^i - \sum_i x_{t-1}^i / e_T^i}{x_{t-1}^{EU}} = \frac{\sum_i \Delta x_t^i / e_T^i}{x_{t-1}^{EU}} \cong \sum_i \frac{x_{t-1}^i / e_T^i}{x_{t-1}^{EU}} \Delta \ln(x_t^i) \tag{1.32}$$

where x_t is the series to be converted from local currency into the common currency, et the exchange rate, i.e. units of local currency per unit of common currency and T the fixed base year.

In two papers on constructing aggregate eurozone data Beyer et al. (2001) suggest the aggregate

$$\Delta \log y_t^{EU} = \sum_{i=1}^{N} \frac{y_{t-1}^{(i)} / e_{t-1}^{(i)}}{y_{t-1}^{EU}} \Delta \log y_t^{(i)}$$

which uses a flexible exchange rate; in this case the weights are $\omega_{it} = \dfrac{y_{t-1}^{(i)} / e_{t-1}^{(i)}}{y_{t-1}^{EU}}$. They construct aggregate growth rates using variable weights, aggregating each country's own growth rate using the corresponding shares of that country in the total aggregate of the previous period as weights. The exchange rate that is used to calculate each country's share is not fixed at a base period but it varies with each period. Using the constructed aggregate growth rates, the level series are obtained by cumulating backwards, starting from a given value at the end of the sample period.

They prefer this method because aggregating levels using a fixed base period leads to distorted aggregate series, as it depends crucially on the choice of base year. Given the substantial relative price-changes due to currency de- and revaluations, using for instance 1979 as base year would lead to aggregate series that are different from the aggregate series constructed using 1999 as base year. Their method has the nice properties that regional and temporal sub-aggregates aggregate correctly and that aggregating a ratio of two variables is equal to taking the ratio of each variable aggregated separately. When aggregating the level series using fixed base period weights, this property holds only for level series and not for growth rate series.

To better compare the method of Beyer et al. (2000, 2001) to the method proposed by Winder (1997), Bosker (2004) writes that

$$\Delta \ln(x_t^{EU}) = \sum_i \frac{x_{t-1}^i / e_{t-1}^i}{x_{t-1}^{EU}} \Delta \ln(x_t^i) \cong \sum_i \frac{x_{t-1}^i / e_{t-1}^i}{x_{t-1}^{EU}} \frac{\Delta x_t^i}{x_{t-1}^i} = \sum_i \frac{\Delta x_t^i / e_{t-1}^i}{x_{t-1}^{EU}} = \frac{\sum_i x_t^i / e_{t-1}^i - \sum_i x_{t-1}^i / e_{t-1}^i}{\sum_i x_{t-1}^i / e_{t-1}^i} \tag{1.33}$$

Comparing the equation above to Winder's, with a fixed base year chosen for the whole sample period, one sees that this fixed base year is each year set at the previous year. Therefore, variable weight growth rate aggregation does take exchange rate effects into account without introducing an unwanted component in the aggregate series. This gives it an advantage over fixed weight level aggregation as this method is unable to properly take exchange rate changes into account and it crucially depends on the chosen base year.

In chapter 2, outlining medium to long run components of the growth rate by generalized dynamic factor model (following the Eurocoin approach), we will observe that the role of common factors in filtering and estimating data is relevant at least as that concerning GDP data to project. In fact, projecting French GDP on European variables (producing a *national indicator*) for example, we produce an indicator very similar to Eurocoin (the medium to long run component of Euro Area GDP growth rate). The Eurocoin (NE) is the projection of the European bandpassed GDP on a set of regressors, that are linear combinations of 157 European variables[23]. These regressors then become an instrument to outline Euro area smoothed growth rate several months ahead of the official GDP estimate. They could also be considered as an indirect and synthetic method of aggregation without using some specific weights.

The Eurocoin approach shows that it is possible to qualify business cycle observations without restricting them to particular countries, as they appear to be regularities common to all decentralised market economies (Lucas 1977).

[23] These time series regard the following European countries: Belgium, Finland, France, Germany, Greece, Italy, Netherlands, Spain, UK.

Annex 1 - The Bry and Boschan Algorithm to Detect the Turning Points

The procedure developed by Gerhard Bry and Charlotte Boschan (1971) is based on the method of the National Bureau of Economic Research. This algorithm is a practical method to compute turning points (TP), particularly with regard to Gross Domestic Product time series. It will be used in chapter 3 to estimate TP in real time and in bandpassed target.

In the quarterly Bry and Boschan (1971) procedure, the following assumptions have to be made:

1. A peak (trough) must be followed by a trough (peak).
2. A cycle (from peak to peak or from trough to trough) must have a duration of at least 5 quarters (or 15 months).
3. A phase (from peak to trough or from trough to peak) must have a duration not lower than 2 quarters (or 5 months).
4. Turning points are not to be situated within the first or last 2 quarters of a data series.
5. The first (last) peak and trough will be higher, respectively lower than values closer to the beginning (end) of the series.

Such assumptions were defined by Bry and Boschan for monthly data; also in this thesis we will use monthly time series (see annex 3) from Thomson Financial Datastream, to estimate Eurocoin and our disaggregated indicators.

The Bry and Boschan procedure is quite structured. At first, the macroeconomic series to be analyzed must be highly smoothed in order to determine the approximate time of the TP and the region where they are likely to be located, and major cyclical swings must be identified. Smoothing is then reduced until the original series is at hand, so that the exact period in which the peaks and the troughs occur can result at the end. All procedures are performed on seasonally adjusted data.

This procedure is particularly suited for data series with many local minima and maxima, as observed in most business cycle indicators concerning aggregated data (e.g. Euro Area GDP) and disaggregated data (e.g., Euro Area sector growth rate, National growth rate, National Household Consumptions).

Considering a quarterly time series $y_t \quad t = 1,...N$ the cycle C_t can be expressed in terms of its TP and these points are the local maxima and minima in the series fluctuations; TP can be defined in Bry and Boschan as follows:

C_t can be located at a peak if $(C_{t-2}, C_{t-1}) < C_t > (C_{t+1}, C_{t+2})$

C_t can be located at a through if $(C_{t-2}, C_{t-1}) > C_t < (C_{t+1}, C_{t+2})$

In this thesis, we apply a Bry-Boschan algorithm to determine the peaks and troughs of a data matrix X, with T monthly time series observations (concerning the period 1995-2010).

For monthly data, the minimum peak-to-trough (trough-to-peak) period is 5 months and peak-to-peak (trough-to-trough) is 15 months. For quarterly p-to-t is 2 quarters and p-to-p is 6 quarters. For annual data, p-to-t is 1 year and p-to-p is 2 years.

More precisely, the Bry and Boschan (1971) procedure for monthly data will consist, essentially, of six steps:

I - In a first step, we have the determination of extreme values and outliers removed and replaced with the Spencer curve value that has, approximately, the flexibility of a five-month moving average but much smoother. The Spencer curve is a complex fifteen-month graduation formula: a weighted moving average with the highest weights in the centre and negative weights at either end. This ensures that curve follows data very closely.

II – In a second step we have the determination of cycles in one-year centred moving average (enforcing alternating peaks and troughs); this second step regards the smoothing of the time series with a moving average and the determination of the approximate TP:

a) identification of points higher (or lower) than 5 months on each side;
b) enforcement of alternation of turns by selecting the highest of multiple peaks (or the lowest of multiple troughs);

III – Determination of peaks and troughs in the Spencer curve:

a) identification of the highest (or lowest) value within ± 5 months of the selected TP in the 12-term moving average;

(b) enforcement of a minimum cycle duration of 15 months by eliminating lower peaks and higher troughs of shorter cycles.

IV – In this step, the smoothing is reduced again and the turning points are readjusted. Peaks and troughs are refined with moving averages determined by the number of months of cyclical dominance (MCD)[24]. For annual data, the cyclical dominance is set to 1 year. Also we enforce alternating peaks and troughs.

V - Peaks and troughs are readjusted in respect to the original time series. We also enforce alternating peaks and troughs, and a minimum p-to-p (t-to-t) period.

VI - In the sixth step TP calculated in the five previous steps are displayed.

An approach to detect TP in the whole Euro area aggregate once they have been detected for each single series (for instance, every disaggregated indicator by sector that we build in chapter 2 and 3 of the thesis), can be carried out by the aggregation procedure proposed by Harding and Pagan (2002). In synthesis, if we have a *K*-dimensional time series of turning points, the

[24] The web-Glossary for Composite Leading Indicator OECD defines the MCD as "the shortest span of months for which the I/C ratio is less than unity. I and C are average month-to-month changes without regard to sign of the irregular and trend-cycle component of the series, respectively. There is a convention that the maximum value of MCD should be 6. For quarterly series, there is an analogous measure, quarters for Cyclical Dominance (QCD), which has a maximum value conventionally defined as 2.

method consists in finding, for each time point *t,* a vector with *K* distances to the nearest peak for every series considered. The median of this vector is the mean distance to the nearest peak for the whole economy. Considering all *t* points, the local minima of this series could be a peak for the whole economy. The same procedure is applied for troughs.

Annex 2 – The Pesaran and Timmermann Test

In chapter 3, to analyze real time performance of the proposed indicators in terms of correct prediction of sign, we use the Pesaran and Timmermann (1992) statistic of directional accuracy (DA). This test analyzes the hypothesis that there is no relationship between the direction of change predicted by a model and the observed change (e.g., in our case, the bandpassed national or sectoral target).

DA test is based on a distribution-free procedure and it is used as a predictive-failure test in the construction of forecasting models (e.g., see Pesaran and Timmermann (1995) and Qi (1999)).

We firstly outline this test in terms of success prediction ratio. Let y_t the target (i.e. the bandpassed growth rate in this thesis) and $\hat{y}_t$ be an estimate in real time of y_t. We assess that $sign(y_t)$ and $sign(\hat{y}_t)$ take value -1 when its argument is negative, and value +1 when its argument is non-negative. When the estimates do not have predictive abilities, the success ratio

$$A_T = \frac{1}{T}\sum_t sign(\hat{y}_t) sign(y_t) \tag{1.34}$$

is very similar to the expected success ratio that would be obtained if y_t and $\hat{y}_t$ were independent.

An estimate of the latter is:

$$B_T = \left(\frac{1}{T}\sum_t sign(\hat{y}_t\right)\left(\frac{1}{T}\sum_t sign(y_t)\right) \tag{1.35}$$

A consistent estimate of $p_y = \Pr\{sign(y_t) = 1\}$ is:

$$\hat{p}_y = \frac{1}{2}\left(1 + \frac{1}{T}\sum_t sign(y_t)\right) \tag{1.36}$$

A consistent estimate of $p_{\hat{y}} = \Pr\left\{sign(\hat{y}_t) = 1\right\}$ is:

$$\hat{p}_{\hat{y}} = \frac{1}{2}\left(1 + \frac{1}{T}\sum_t sign(\hat{y}_t)\right)$$

Under independence of $\hat{y}_t$ and y_t we have:

$$DA \equiv \frac{A_T - B_T}{\sqrt{V_{DA}}} \xrightarrow{d} N(0,1) \tag{1.37}$$

where the variance of A_T and B_T under the null hypothesis and by using the estimate (1.36) is:

$$V_{DA} = 16\frac{T-1}{T^2}\hat{p}_{\hat{y}}(1-\hat{p}_{\hat{y}})\hat{p}_y(1-\hat{p}_y) \tag{1.38}$$

- ***The test in the 2 x 2 case: another representation***

Supposing, in this sub-section, the case with 2 categories into which the realizations of y and x fall and we follow, in detail, the original article of Pesaran-Timmermann (PT). Let $x_t = \hat{E}(y_t|\Omega_{t-1})$ as the predictor of y_t that is formed with regard to the information available at time $t-1, \Omega_{t-1}$. Supposing n is observations on (y_t, x_t). The PT test is based on the proportion $\hat{P}$ of times that the direction of the change in y_t is correctly "predicted" (i.e. "nowcasted" with regard to this thesis).

Let the following assumptions:

$Y_t = 1$ if $y_t > 0$,

$Y_t = 0$ otherwise,

$X_t = 1$ if $x_t > 0$,

$X_t = 0$ otherwise,

$Z_t = y_t x_t = 1$ if $y_t x_t > 0$,

$Z_t = 0$ otherwise.

Let also $P_x = \Pr(x_t > 0)$ and $P_y = \Pr(y_t > 0)$, then $\hat{P} = n^{-1}\sum_{t=1}^{n} Z_t = \overline{Z}$

On the assumption that y_t and x_t are independently distributed and therefore x_t has no power in predicting y_t, $n\hat{P}$ has a binomial distribution, and

$mean\ (n\hat{P}) = nP_*$

with $P_* = \Pr(Z_t = 1) = \Pr(y_t > 0,\ x_t > 0) + \Pr(y_t < 0, x_t < 0) = P_y P_x + (1 - P_y)(1 - P_x)$

In the case in which P_x and P_y are known, this test is based (as in the equation (1.37)) on the standardized binomial variate:

$$S_n = \left\{ \frac{P_*(1 - P_*)}{n} \right\}^{-\frac{1}{2}} (\hat{P} - P_*) \qquad (1.39)$$

that is asymptotically distributed as . $N(0,1)$

The S_n test can also be extended to situations in which the information on Y_t and X_t are categorized into more than two classes.

Chapter 2 – Empirical Applications Disaggregating Eurocoin: Building the Models

In chapter 1, New Eurocoin (NE) indicator is described through the projection of c_t (the whole Euro bandpassed gross domestic product) on a set of regressors – the linear combination of variables contained in the *Thomson Financial Datastream* used by the Bank of Italy. NE provides a summary index of the medium to long run component of GDP, but only for the whole Euro Area aggregate. Therefore, NE does not produce a monthly estimate by country, sector or GDP components.

In fact, NE provides an index of the current economic situation in the Euro Area, extracting from the Data Source relevant information which represent the main sources of variation, in order to track the entire underlying GDP for the whole Euro area, thus it is an "*aggregate indicator*".

The main aim of this section is to propose a new theoretical framework for disaggregated business cycle analysis *by sectors, countries and expenditure components* based on the Eurocoin approach that we have outlined in section 1.3.

Objectives developed in this chapter are:

a) in section 2.1, we describe the dataset and the data treatment used in building our disaggregated estimates and there is an explanation why we decide to construct some disaggregated smoothed estimates of GDP with the Eurocoin approach;

b) using the Eurocoin approach for building European sectoral estimates of the medium to long run component of GDP growth **(MLRG)** (section 2.2). Main macro sectors analyzed: Manufacturing, Energy, Construction, Financial, Trade-Transport-Communication.

c) We build some *national smoothed indicators* based on a generalized dynamic factor model, by projection national GDP on European or national common factors. The following countries are investigated: Belgium, Italy, France, Germany, Netherlands, Spain (section 2.3).

d) Developing Eurocoin methodology by dividing European variables used to build common factors, in real and financial variables, we analyze their impact on smoothed GDP. This subdivisions among real economy and financial economy is substantially confirmed in Forni et al. (2003), where they study the impact of financial variables on real data. We show that a combination between "real MLRG" and "financial MLRG", obtained projecting Euro Area GDP respectively on real and financial variables, can compete with truncated band-pass filter within the sample in terms of correlation and RMSFE (par.2.4).

e) Finally, we estimate smoothed GDP components (Household Consumptions, Investments, Imports and Exports) and their volatility, using the Eurocoin methodology to calculate these variables (par 2.5).

We analyze the "medium to long-run growth" that is not precisely the growth-rate cycle or the "business cycle", as in the definition of a cycle even the oscillations of a period longer than 8 years are generally removed (for different definitions of the cycle, see Stock and Watson, 1999).

In fact, we are interested in the real time forecasting performance of our indicators with respect to a measure of the "trend-cycle GDP growth" obtained in the middle of the sample by a band pass bilateral filter on GDP growth components.

The innovation of this research are some disaggregated procedures, based on the Eurocoin methodology in order to obtain smoothing of a stationary time series, therefore avoiding the occurrence of end-of-sample deterioration. The objective is to build sectoral and national real-time monthly estimates of GDP growth purified from seasonal and other short run fluctuations, as well as from errors in the measurement of GDP, and highly reliable at the end of the sample. Concerning local and sector-specific shocks, contrary to the Eurocoin, we maintain those useful for the description of disaggregated components of Euro growth:

- by projecting bandpassed sectoral growth on European common factors of generalized dynamic factor models that we use in our applications. In this case, the common factors are the same factors used in building the Eurocoin index;
- by projecting the smoothed national or European growth on national factors (only built with the macroeconomic variables regarding a specific country) and/or European factors;
- by the projection of expenditure components of the growth (e.g. Household Consumptions) on the European common factors.

Therefore, we build two types of “common factors”:

- “European factors” that are a linear combination of the 157 variables (belonging to different Euro Area countries);
- “National factors” that are a linear combination of the national variables included in our Dataset (e.g., for France we have 23 macroeconomic variables that are used to construct the set of regressors useful for the estimation of past and future MLRG values) concerning a specific European country.

In this chapter, we build *ex post (in-sample)* estimations that are therefore generated for the same dataset that was used to estimate the parameters of the model. For this reason, it is expected that these forecasts are relatively good. We will compare these estimates by using descriptive statistics to bandpassed components of GDP and to Eurocoin indicator (we will briefly outline the bandpassed target in section 2.1 and more deeply in chapter 3 for a real time estimation at the end of the sample).

On the other hand, *real time performance* is a reasonable approach for the examination of estimate accuracy. Not all observations can be used in estimating parameters. The latter sample will be used to build pseudo estimates by a recursive or a rolling window (see section 3.1). In chapter 3, we will test our models using *pseudo real time estimates* at the end of sample. Real time estimates will be compared to bandpassed Euro Area growth components and we will assess if the in- sample results are certifiable.

The contribution of this chapter is that of building some comprehensive measures for each economic activity and European country, like GDP, that show the underlying direction in which disaggregated components of the European economy move free from problems listed below. In fact, Euro Area GDP is a comprehensive measure of economic activity, but:

1. it is released on a quarterly frequency, with a certain delay and may be subject to significant revisions afterwards.

2. GDP growth may be high or low in any quarter depending on seasonal effects and measurement errors. Then GDP can be misleading;

3. it could be influenced by factors affecting only a particular country. These factors are not important for outlining the health of the euro area economy as a whole, but they can be used to asses the national cycle of this country. We will propose this approach in the section 2.3 concerning the "National Eurocoin"

Removing erratic components can also be done, for example, by applying a band pass filter to the GDP growth series. This technique, however, presents the same problems in terms of frequency and timeliness, producing some estimates that deteriorate at the end of the sample.

This chapter is dedicated to the ex-post analysis of business cycle in the Euro Area compared to a bandpassed estimation of our data, that is our empirical target. It is organized as follows:

- Paragraph 2.1 presents the dataset used in the empirical application and treatment of data
- par. 2.2 presents a "Sectoral Eurocoin (SE), therefore an application of the Eurocoin approach by sectors
- in par 2.3 we build a "National Eurocoin", i.e. we construct some national indicators for the smoothed GDP in Euro Area
- in par 2.4 we show the impact of real and financial variables on nowcasting the medium to long run component of the growth
- in 2.5 we present some indicators to estimate the smoothed growth rate for each components of GDP (Consumptions; Foreign Trade; Investments).

2.1 Introduction: Dataset and Data Treatment

Nowcasting GDP requires to focus on times series data that can provide information on the current state of the economy. There are numerous macroeconomic time series with shorter publication delays than GDP. This is mostly the case for monthly statistics related to employment, industrial production, financial variables or business surveys published by Central Banks or National Statistics Institutes.

At this stage, two main approaches exist in econometric literature to choose the variables useful for the nowcast of growth rate; the first focus on a limited number of series and it consists in selecting a reduced number of variables and tracking their development. The selection criteria are generally based on:

- the ex-post ability of the series to reproduce reference time series movements;
- a priori belief based on economic theory;
- the choice can almost be judged as subjective.

The series can either be individually tracked or aggregated in a synthetic index. The former is a strategy that has been adopted by the Conference Board and the OECD.

In the present paper, we will be using generalized dynamic factor models, explained in par. 1.3, to construct some monthly indicators of economic activities (by sector, country and GDP components) in Euro Area, and we assume that a better representation for the economic development could be obtained by considering a large dataset of European variables. However, we show that a more parsimonious use of the variables based on the "nationality criteria" (by a smaller dataset of national data only regarding a specific country) can be useful to analyze National Household Consumptions (e.g. see the Italian case in section 3.5). Our objective here is how to synthesize information contained in a large sample of macroeconomic and financial time series, rather than how to select individual series. Data and variables used to build the common factors are presented in table 2.1. They belong to the *Thomson Financial Datastream* used in our research (a detailed description of these data is given in the Annex to this chapter). As per series of gross domestic product by sector and country, they are collected from Eurostat and ECB Statistical Dataset available on their official site.

We dispose of a dataset consisting of 157 monthly macroeconomic variables during the period between January 1987 and March 2011. The main blocks of macroeconomic indicators are as follows (see table 2.1):

- Business and consumer confidence indicators – the largest block;
- Industrial production indices;
- OECD Composite Leading Indicator;
- Producer price index for: intermediate and capital goods; energy, industry, investment and intermediate goods; durable and non durable goods;
- Retail Sales;

- Variables describing external transactions: exports and imports of goods and services.
- Financial data: monetary variables, interest rates, effective exchange rates.

Table 2.1: Variables used in Estimation by Data Source

Data Source	***Variables***
Surveys	31
Leading Indicators	6
Demand Indicators	12
Industrial Production	32
Wages Indicators	2
Employment Indicators	5
Producer Price Index	26
Exchange rates	3
Imports-Exports	8
Money Supply	8
STANDARD & POOR'S INDEX	7
(Italy, Germany, USA, UK) SPREAD	10
Benchmark Bond	7
TOTAL	***157***

We focus on medium to long-run components of the growth (MLRG), i.e. the smoothed components of GDP growth rate obtained by removing the fluctuations of period shorter than or equal to one year, and it bears no relationship to any definition of trend.
Our elaborations are based, in this research, on the GDP expressed in constant prices.

The main preliminary steps followed to build our indicators are as follows:

a) The first step in building our indicators is to apply statistical and econometric theories to stationary data: this means that the indicator will track changes in GDP, and not its level.

All 157 series are transformed to remove seasonal factors and non-stationarity. Seasonal adjustment is conducted by regressing variables on a set of seasonal dummies, while non-stationarity is removed by first differencing or log-differencing.
b) According to current practice in conjunctural analysis, our dataset is organized to include groups of variables that are *leading, lagging* and *coincident* with respect to the GDP.

Assuming variables in the dataset are x_{it} and that $x_{it} = u_{t-k_i} + \xi_{it}$.

Let $i = 1,2,...n$ and k_i take the values 0,1 or 2. We then have three subgroups:

- the lagging variables (those loading u_{t-2} with $k_i = 2)$;
- the coincident or central variables (those loading u_{t-1} with $k_i = 1)$;
- the leading variables (those loading u_t with $k_i = 0)$: they are crucial to obtain a good estimate of the medium to long run component of GDP (or of its components) at the end of the sample containing information on future GDP values.

Monthly indicators are commonly used in the prediction of current data on GDP before the data are available. For the Euro area, a flash estimate of GDP is released by Eurostat about six weeks after the end of the reference quarter, and a full set of indicators for the second quarter of the year is not available any earlier than the GDP flash estimate.

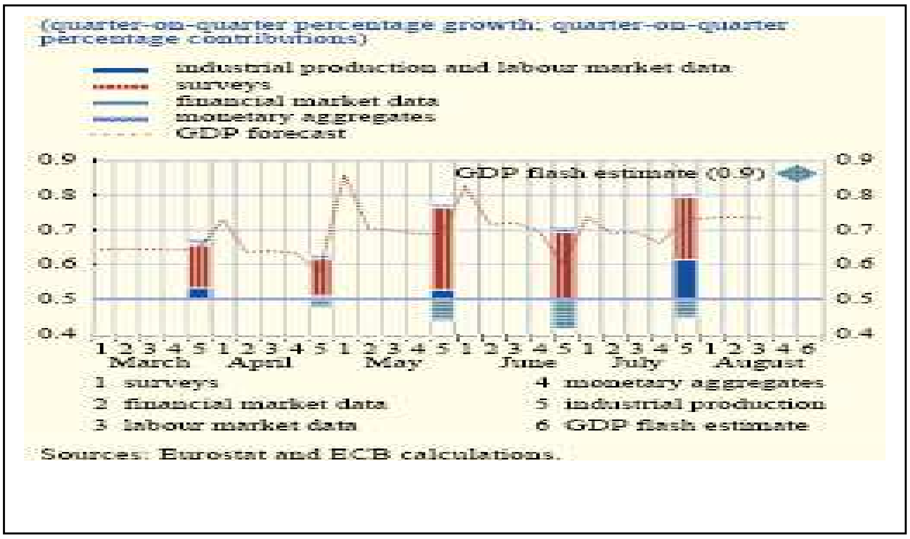
For instance, the GDP flash estimate for the second quarter is available in mid-August, and since financial market data are published on a daily basis, a monthly average of these data is available at the end of the month. Consumer and business surveys are available at the end of the reference month, while data concerning monetary aggregates and unemployment rate are available with a delay of about one month. Finally, industrial production data are published about six weeks after the end of the reference month.

Thanks to their earlier availability, financial market data and survey are important sources of information at longer forecasting horizons, while industrial production data gain importance mainly for the (final) forecasts conducted one month ahead of the release of GDP data, when they become available for the quarter of reference.

Most timely variables (such as Purchasing Managers Indexes, Consumer Surveys, Business Climate Indexes, etc.) are generally not comprehensive and smooth. Other series, such as Industrial Production and Exports, ignore large portions of economic activities. Furthermore, they all exhibit heavy short-run fluctuations and could provide contradictory signals.

In the real chart below, it is possible to see, in a dynamic factor model (*European Central Bank analysis – ECB monthly bulletin 04/2008*), the different contributions of data groups to the GDP forecast for the second quarter of 2006. Analysis of information obtained from individual data groups may occasionally provide conflicting signals on the whole economic activity. In fact, in figure 2.1, after predicting values slightly above 0.6% throughout March, the GDP forecast jumped to above 0.8% with the release of survey data for April. It fell back thereafter, once other data came in. However, survey continued to give positive signals. Final forecasts from mid-July remained above 0.7%, thanks to a positive contribution from industrial production.

Figure 2.1: Contribution of data group to the GDP forecast in a dynamic factor model



Source: ECB Monthly Bulletin – April 2008

Note: the contribution is expressed in terms of deviation from average GDP growth

c) A key step in building the indicator has been to isolate the common components by decomposing each variable into two uncorrelated components: a "common component" and an "idiosyncratic component" , to remove Measurement Error, Local and Sectoral Shocks. The idiosyncratic component captures the effect of shocks affecting only that variable, while the common component depends on a small number of common shocks, which affect all the variables with different weights and lags. In fact, the coincident indicators as the ones we build in this chapter, are formally defined as common components of GDP growth rate, after this variable is filtered to eliminate high frequency and short-run variations (i.e. at frequencies of less than fourteen months), as well as seasonal movements, in order to reveal the underlying medium and long-run tendency of the economy.

Previous research has relied on "two-sided filters" to eliminate seasonal and short-run high frequency noise (see Baxter and King, 1999, Christiano and Fitzgerald, 2003). These filters perform well in the middle of the sample, but they work badly at the beginning and at the end of the sample, since they require knowledge of the future values of GDP, which of course we do not have.

This is the main technical reason why it is worthwhile to develop the disaggregated indicators that we will present in detail in the following sections.

2.2 Building a Sectoral Eurocoin

New Eurocoin indicator, published monthly by the Banca d'Italia and CEPR, provides a summary index of the medium to long-run component (MLRG) of the entire GDP for the whole aggregate Euro area. The innovation of this section are some procedures based on Eurocoin methodology to estimate sectoral MLRG concerning Euro Area: performance of these indicators is assessed with respect to a measure of the "trend-cycle GDP growth".

In sub-section 2.2.1, we explain reasons why sectoral indicators are built, the approach to construct them and compare interpolated sectoral growth with bandpass filter. In 2.2.2 and 2.2.3 we also analyze ex post estimate and their relation with Eurocoin indicator. Therefore, we do not carry out, in this chapter, a real time estimation.

2.2.1 Why a Sectoral Eurocoin?

A *sectoral estimate* of GDP is released by Eurostat about ten weeks after the end of the reference quarter.

"As errors in Flash estimates tend to be relatively large, National Statistical Institutes tend to release them at higher level of aggregation than preliminary Quarterly National Accounts estimates" [25]. In order to have more disaggregated and timely data, we develop a "Sectoral Eurocoin" (that we name SE in the following).

In this chapter, we also develop ex post estimates being generated for the same dataset that was used to estimate the parameters of the model.

Then, we produce a monthly estimate of growth purified from erratic components, following sectoral breakdown of quarterly gross value added available from Eurostat and European Central Bank. Sectoral composition influences the characteristics of a business cycle, such as its length and amplitude. Factor model is used to meet the macroeconomic behaviour on the basis of disaggregated data (the sectors).

In this section a sectoral version of the Eurocoin indicator is proposed as today, Eurocoin is only used to derive aggregate European growth.

Using Eurocoin methodology, in this paragraph we will develop a disaggregated analysis of the medium to long run components of GDP (MLRG) in the European economy, studying interrelations and characteristics of sectoral growth rates.

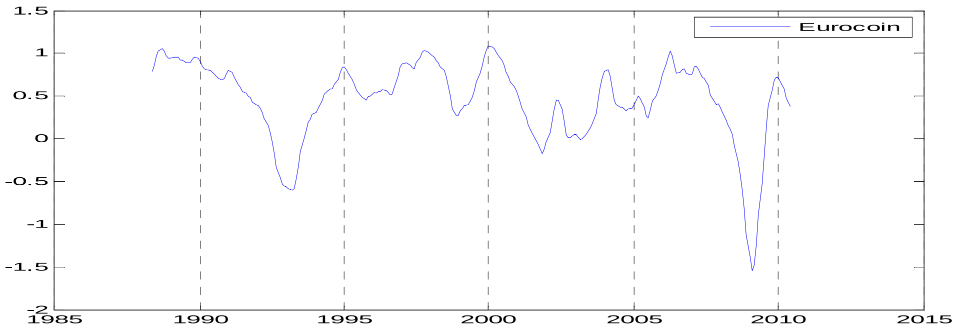
We also carry out careful investigation on sectoral volatility in the Euro Area as its impact on GDP it is not entirely clear. "Industry seems to be an important factor in shaping the aggregate business cycle, other sectors may contribute as well, both directly through a particular sectoral volatility or indirectly through sectoral linkages". However, some sectors, "despite having the

[25] "Producing flash estimate of GDP", Note prepared by the *United Nations Economic Commission for Europe*, 2003.

highest relative volatility per country, contribute little to the EU-wide volatility. This is possibly correlated to the fact that these sectors are relatively more exposed to country–specific fluctuations" (*Source*: Sectoral Specialisation in the EU – Occasional paper series NO.19/2004- European Central Bank).

For ECB (European central bank) statistics, "Industry branch" includes data concerning the following sectors: Manufacturing, Energy, Constructions; while "Service branch" includes data concerning Trade, Hotel and Restaurants, Transport and Communications, Financial.

Figure 2.2: Eurocoin as calculated by the Bank of Italy



The steps followed to obtain these indicators include the projection of sectoral added values on European factors (that are the combinations of the 157 variables contained in the Thomson Financial Datastream and used by the Bank of Italy to build Eurocoin): the resulting monthly indexes are a "smooth version" of sectoral growth rate, that give an early estimate of the Euro area growth performance in terms of quarter-on-quarter changes in GDP.

By the Eurocoin approach, monthly estimation concerning sectoral medium to long-run growth $\hat{c}^{s}_{T}$ will be obtained projecting sectoral bandpassed GDP on common smooth factors (see chapter 1, par 1.3):

$$\hat{c}^{s}_{T} = A^{s}{}_{1F} F_{1t} + A^{s}{}_{2F} F_{2t} + ... + A^{s}{}_{mF} F_{mT}$$

We smooth sectoral stationary series by using only current (contemporaneous) values of a large dataset so that, contrary to the bandpass filters, no end-of-sample deterioration occurs.

In this chapter, we will compare our sectoral ex post oral indicators with bandpassed sectoral growths (our empirical target) that is built using a bandpass filter (see section 3.2.1 for explications concerning the building of the target).

In chapter 3, we will test our sectoral models in real time. Bandpassed sectoral growths produced in this chapter will be compared to pseudo real time estimates. In addition, we will assess if the in-sample results can be confirmed.

According to Altissimo et al. (2006), within a finite sample, the following approximation of the target can be obtained by augmenting y_t^s with its sample mean $\hat{\mu}$ in both infinite directions:

$$c_t^{*s} = \beta(L) y_t^{*s}, \text{ where } y_t^{*s} = \begin{cases} y_t & if \quad 1 \le t \le T \\ \hat{\mu} & if \quad t < 1 \ or \ t > T \end{cases} \qquad (2.1)$$

Since y_t, the sectoral growth rate, is observed only quarterly, while we are interested in a monthly indicator of economic activity, in figures 2.3 a simple interpolation is chosen to calculate the two missing points for each quarter, assuming that y_t remains unchanging within a quarter.

These filters work well in the middle of the sample, but they perform badly at the beginning and at the end of the sample, since it is worthwhile to develop the disaggregated indicators that we analyze more deeply in sub-sections 2.2.2-2.2.3.

Figures 2.3 below show approximate sectoral MLRG (our *empirical target*), that is obtained by removing, from the quarterly growth of real GDP, any fluctuations of a period shorter than or equal to one year, that is compared to quarterly growth rate for the following sectors:

- Manufacturing, Energy;
- Construction;
- Trade, Transport, Communication;
- Financial;
- The whole "Industry" aggregate;
- The whole "Service" aggregate.

Figure 2.3/A** - c_t^* **and monthly quarter-on-quarter Manufacturing growth rate

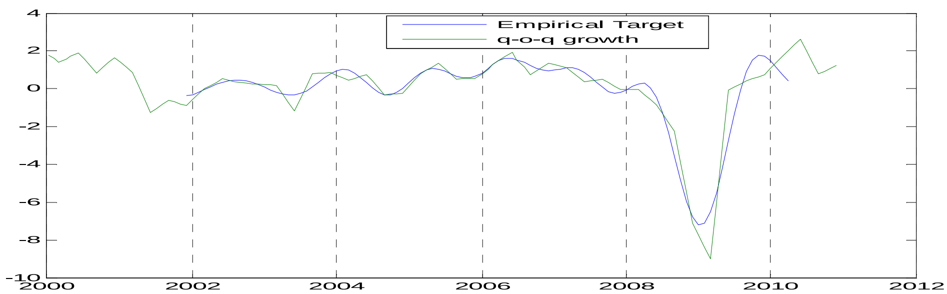


Figure 2.3/B - C_t^* and monthly quarter-on-quarter Manufacturing plus Energy growth rate

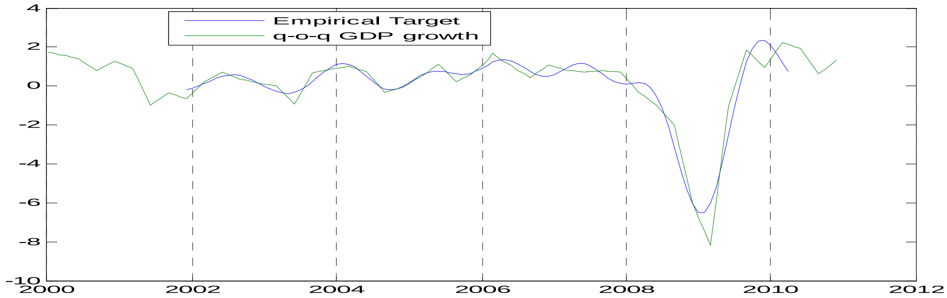


Figure 2.3/C: C_t^* and monthly quarter-on-quarter Construction growth rate

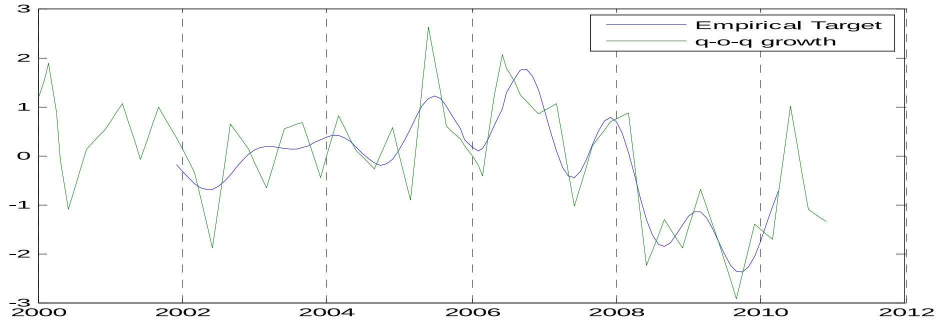


Figure 2.3/D - C_t^* and monthly quarter-on-quarter Trade-Transport-Communication growth rate

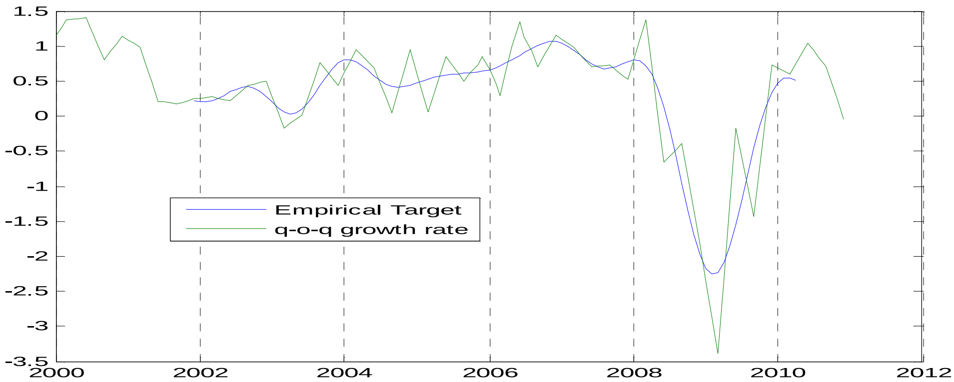
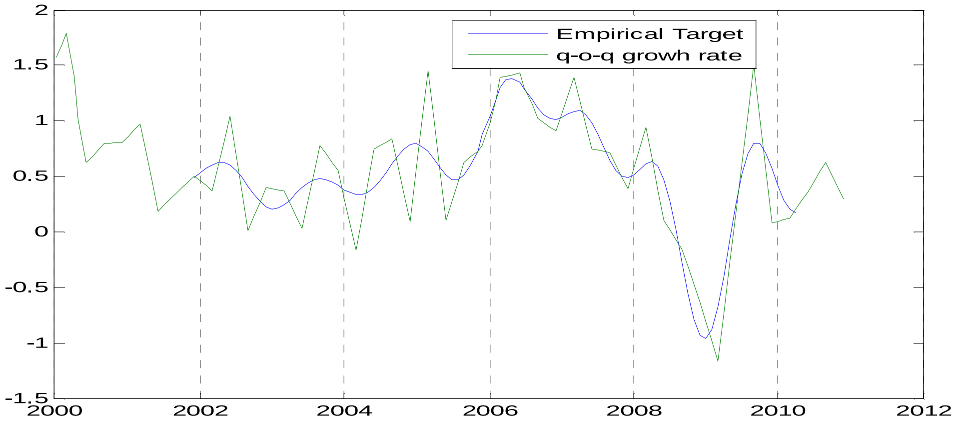


Figure 2.3/E **-** C_t^* ***and monthly quarter-on-quarter Financial growth rate***



Concerning the figures above and each sector analyzed, we observe that the main difference between MLRG and GDP growth is that the former, free from short-run volatility, is far smoother, so that it shows the underlying growth of the economy more clearly. Therefore, a downturn (upturn) is always followed by several months of increasing (decreasing) monthly growth rate. This is particularly evident for Construction, Trade and Financial.

In figures 2.3/F and 2.3/G, we show approximate sectoral MLRG (our *empirical target*) versus quarterly growth rate, by aggregating the following series:

- Manufacturing, Energy and Construction data to produce an “Industry” MLRG;
- Trade-Transport-Communication-Financial series to produce a “Service” MLRG (medium to long run growth rate).

Figure 2.3/F : C_t^* ***and monthly quarter-on-quarter Industry growth rate***

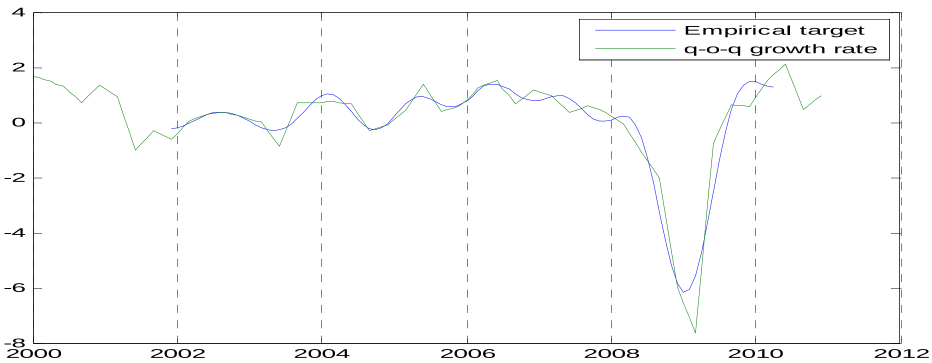


Figure 2.3/G : C_t^* ***and monthly quarter-on-quarter Service growth rate***

Bandpassed filters work well in the middle of the sample, but they perform badly at the beginning and at the end of the sample, since it is worthwhile to develop the disaggregated indicators that are analyzed in detail in subsequent subsections.

2.2.2 Ex-Post Estimation of Sectoral Eurocoin versus Bandpassed Growth

In this section and in the next, we use generalized dynamic factor model to develop sectoral Ex post MLRG.

Ex post MLRG are compared to data that are bandpassed using equation (2.1). In the following tables we show the main descriptive statistics from 1995 to 2010 for bandpassed industrial and service sectors, disaggregated by their main components.

We apply the final estimates of the sectoral growth rate as reported in GDP "vintage" made available in February 2011: data concerning sectoral breakdown of quarterly gross value added are available in Eurostat Statistics Database by themes and European Central Bank Monthly Bulletin (Euro Area Statistics on line). The same holds true for all other monthly series, as vintages we use for most of the monthly macroeconomic variables in the Thomson Financial Datastream are not available.

For this reason, ex post estimate is looked at in this section by analyzing the in-sample 1995-2002. In chapter 3, the period 2003-2010 will be analyzed in real time with the end of the sample.

Table 2.2/A Correlations among sectoral bandpassed growth rates in 1995-2010

	EUROCOIN	INDUSTRY	CONSTRUCTIONS	FINANCIAL	TRADE
EUROCOIN	1	0.94	0.53	0.87	0.94
INDUSTRY	0.94	1	0.33	0.74	0.88
CONSTRUCTION	0.53	0.33	1	0,82	0.54
FINANCIAL	0.87	0.74	0.44	1	0.74
TRADE	0.94	0.88	0.54	0.74	1

We detect, by ex post estimate, that Industrial (Manufacturing + Energy) and Trade-Transport sectors are those with a medium to long run component of GDP most correlated with Eurocoin indicator. All sectors considered in this table are highly correlated with Financial component (that also includes real estates, renting and business activities).

With regard to the value added by economic sector, we remark that Industry actually represents about 25 % of the whole GDP; Services about 50%; Public Administration 25%.

Table 2.2/B Sectoral statistics concerning bandpassed MLRG in 1995-2010.

	MONTHLY MEAN (1995-2008)	**MONTHLY MEAN (1995-2010)**	**MONTHLY MEAN (2008-2010)**	**VARIANCE (1995-2008)**	**VARIANCE (1995-2010)**
EUROCOIN	**0.55**	**0.41**	**-0.45**	**0.08**	**0.31**
MANUFACTURING PLUS ENERGY	**0.54**	**0.28**	**-0.84**	**0.34**	**2.07**
MANUFACTURING	**0.60**	**0.30**	**-1.02**	**0.56**	**2.58**
CONSTRUCTION	**0.21**	**- 0.05**	**- 1.18**	**0.64**	**0.92**
TRADE-TRANSPORT	**0.68**	**0.50**	**- 0.29**	**0.13**	**0.47**
FINANCIAL	**0.75**	**0.64**	**0.18**	**0.10**	**0.18**

In sample (ex-post) estimates concerning sectoral analysis of medium to long run growth rates, following sectoral models that we have developed (and that will be tested in real time in chapter 3) show that:

- during the period 1995-2010, higher mean concerning quarter-on-quarter monthly changes in GDP are those connected to Trade-Transport-Financial sectors. During the 2008-2010 economic crisis, the only sector with a non-negative growth mean is the Financial;
- 1995-2008, the lower sectoral growth means are those concerning the Construction and Manufacturing-Energy sectors. During the 2008-2010 economic crisis, the sector with the most negative growth mean is Construction;

- 1995-2010, higher sectoral volatility concerns Manufacturing-Energy and Construction sectors. During the 2008-2010 economic crisis, sectoral volatility increased significantly in all sectors (with a maximum for Manufacturing), while it increased slightly in the Financial sector;
- looking at aggregated data produced by classic Eurocoin, we observe that this non-sectoral indicator shows quite a slight variance compared to disaggregated indicators (Manufacturing, Energy, Trade, Transport) different from the Financial sector.

In this sub-section, we compare bandpassed growth by sector (our empirical target) with sectoral ex post estimation with regard to the indicators that we build by using the Eurocoin approach. We take $c_t^{*s}(T)$ as our target (see equation 2.1); therefore, we analyze the performance of the "Sectoral Eurocoin" $\hat{c}_t^{\,s}$ at time T (built using generalized dynamic factor model), with $t \leq T-12$, by the difference between our indicator at time t and the approximate target at t that are obtained using data up to T.

Monthly length of the sample 1995-2002 is equal to T=90. For the period T-78 ≤ t ≤ T-13, we are interested in the ability of the ex post indicator to approximate the target as measured by the root mean-square error

$$\sqrt{\frac{\sum_{t=T-96}^{T-13}\left[\hat{c}_t^{\,s}(T)-c_t^{*s}(T)\right]^2}{66}} \tag{2.2}$$

The results of this statistic will be indicated in the tables of this chapter as " *Rmse with respect to c**".

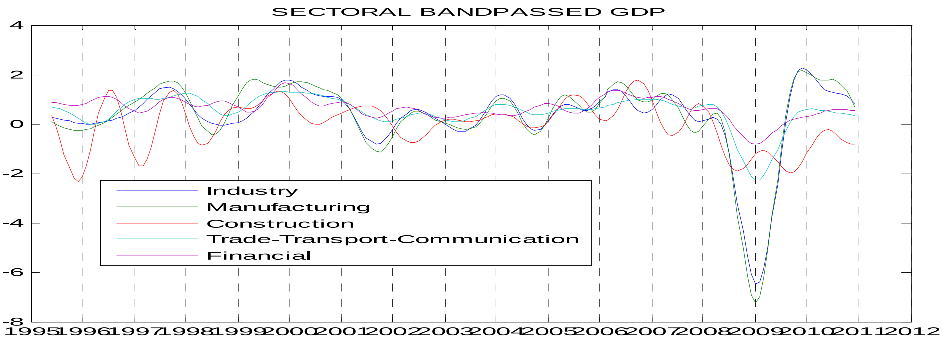
Table 2.2/C – RMSFE among Ex post Estimate and Sectoral bandpassed growth

SECTORS	**1995,6,2002,12** ***Rmse with respect to c*".***
Manufacturing + Energy	**0.25**
Manufacturing	**0.47**
Construction	**0.76**
Trade-Transport-Communication	**0.18**
Financial	**0.19**

Table 2.2/C above shows real time forecasting performance of the in-sample indicator with respect to a measure of the "trend-cycle GDP growth" obtained in the middle of the sample by a band pass bilateral filter on GDP growth. Regarding the nowcast of each sectoral target c_t^*,

we see that the aggregates Financial, Trade-Transport-Communication, Manufacturing + Energy score remarkably higher than Manufacturing and Construction. Moreover, Financial sector in sample indicator is able to approximate fairly well its target.

Figure 2.3/H - Sectoral Bandpassed Growth



2.2.3 Growth Cycle Phases and main Figures Concerning Sectoral Analysis

The figures below (from 2.4 to 2.8) show the outlook for economic activities, using Eurocoin methodology by **ex post estimates.** In addition to rigorous statistical analysis developed in sub-sections 2.2.1 and 2.2.2 for the in-sample period 1995-2002, these figures can be useful in the graphic analysis of the relations among different sectoral smoothed growths within a large period of time .

Following OECD terminology used to specify growth cycle outlook by CLI (Composite Leading Indicators)[26], we define the following cycle phases as:

- Expansion (increase with a growth higher than 0%);
- Downturn (decrease with a growth rate above 0%);
- Slowdown (decrease with a growth rate below 0%);
- Recovery (increase below 0%).

In figure 2.7 below we observe different sectors simultaneously; in June 2010 (our last estimates, while dataset data are available until 2010 Q1) it was possible to determine this outlook for Euro Area:

- Manufacturing: Downturn
- Financial and Trade: Downturn.

Following ECB breakdown of gross domestic product, we consider that "Industry" business cycle include the following sectors:

- Manufacturing;
- Energy;

[26] See Annex to this chapter

- Constructions.

“Service” branch include the following sectors:

- Trade;
- Transports;
- Communications;
- Financial.

Figure 2.4: Medium to long run component (MLRG) of the Euro Area Manufacturing growth

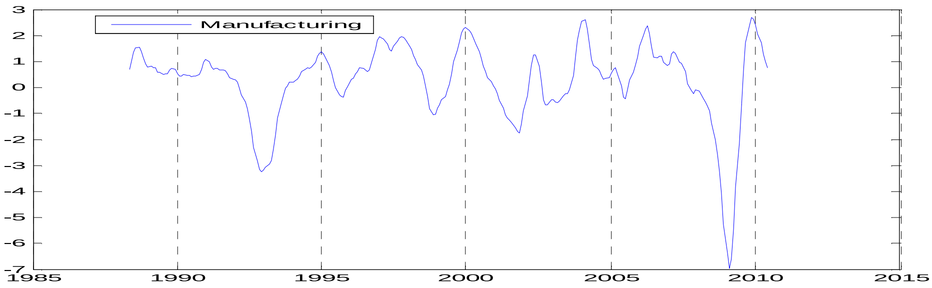


Figure 2.5: MLRG of the Euro Area Trade-Transports-Communications

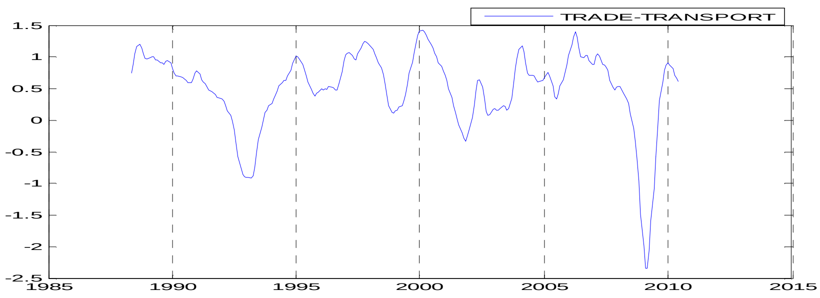


Figure 2.6: MLRG of the Euro Area Financial sector

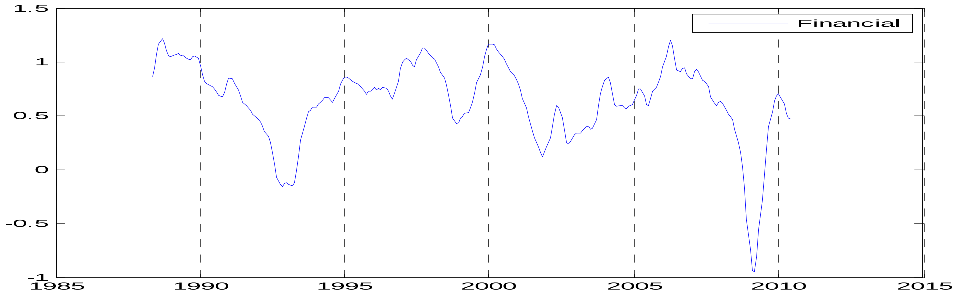
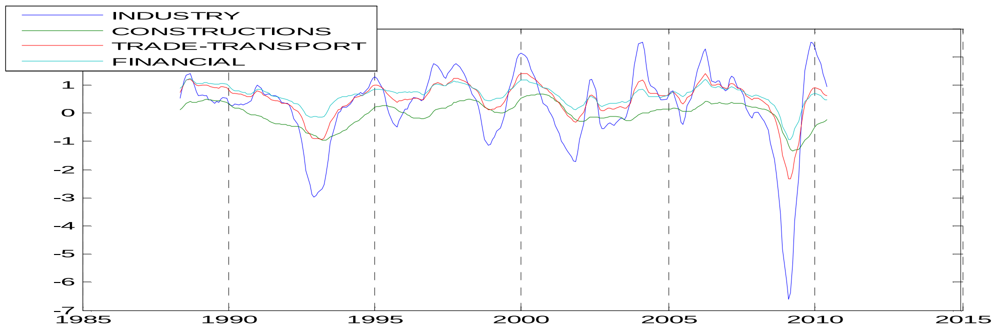
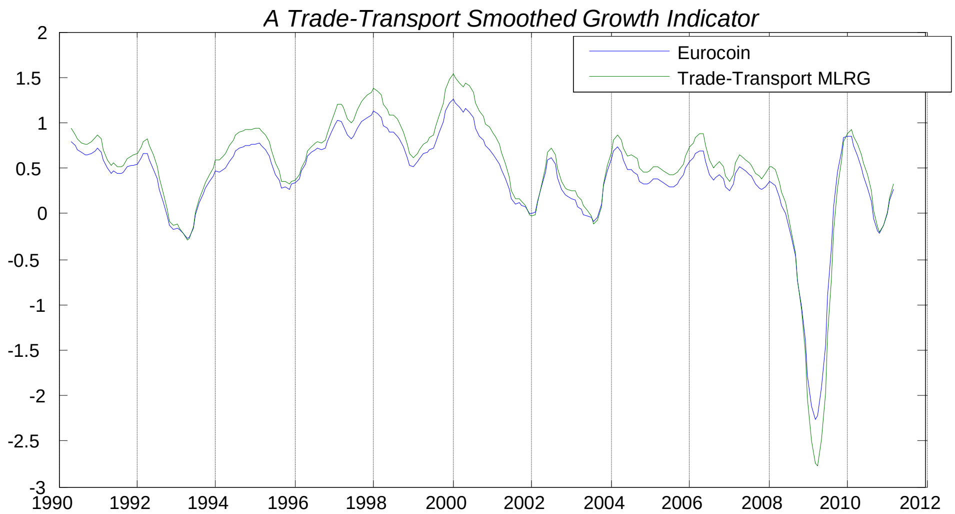


Figure 2.7: Sectoral MLRG in Euro Area



In figure 2.8/A, we compare the ex post Trade-Transport–Communication MLRG with Eurocoin indicator; we find that it is a good proxy of Euro Area GDP, how it also is indicated by tables 2.2.

Figure 2.8/A



In figure 2.8/B-C, we outline *Ex post - medium to long run component* of Industry (aggregating data concerning Manufacturing and Constructions) and Services sector growth rate (aggregating data concerning Trade-Communications and Financial sector). As we show in figure 2.8/B, MLRG concerning *Services* Sector (obtained by aggregation of data concerning Trade, Transports, Communications and Financial) is very similar to Eurocoin indicator. We observe the following correlations :

Table 2.3

Eurocoin versus Services	0.98
Eurocoin versus Industry	0.93
Services versus Industry	0.94

Figures 2.8/B: MLRG of the Euro Area: Industry and Services versus Eurocoin

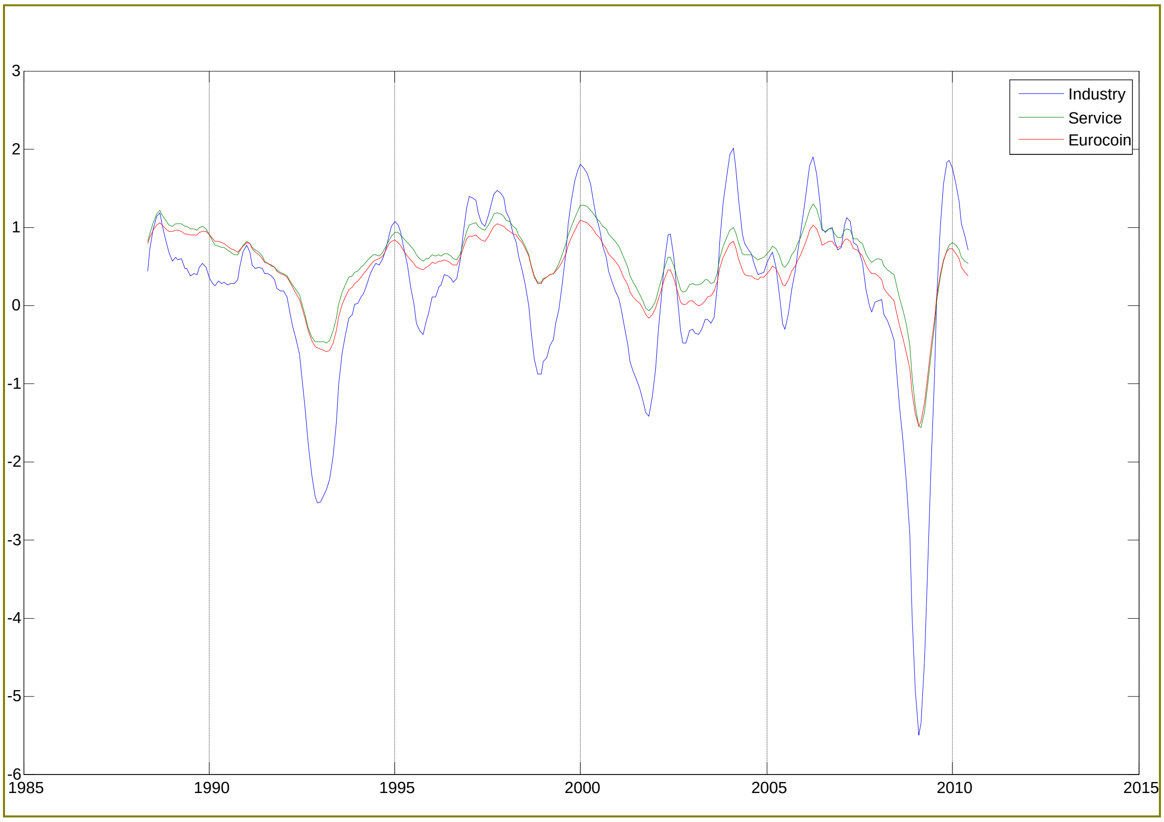
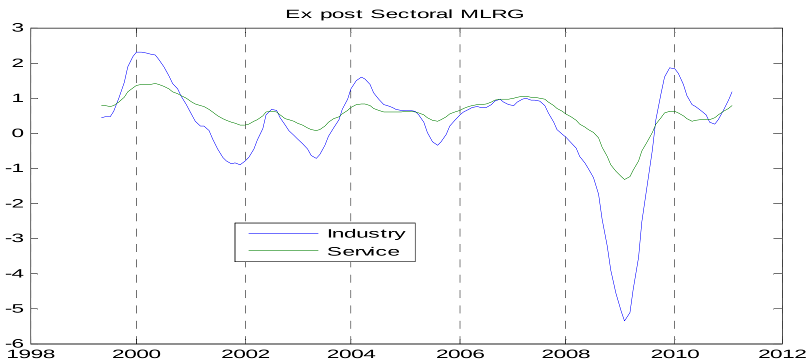


Figure 2.8/C



Using Eurocoin methodology that we have outlined in chapter 1, in this section we have built some disaggregated models for the analysis of the medium to long run component of GDP (MLRG) in the European economy, studying interrelations and characteristics of sectoral growth rates. The main macro sectors analyzed are: Manufacturing, Energy, Construction, Financial, Trade-Transport-Communication.

2.3 Building National Business Cycles

After having analyzed disaggregated sectoral performance in 2.2, in this section, using Eurocoin methodology, we show some approaches and results to obtain the smoothing of a stationary time series so as to avoid the occurrence of highly unreliable estimates at the end-of-sample. The objective is to build national real-time monthly estimates of GDP growth, purified from erratic components (short-run volatility, measurement errors, seasonal variations).

Our main goal is to build a *national Eurocoin*, by the projection of national and Euro Area GDP on European and/or national factors. Our exercise will consider the following countries: Belgium, Italy, France, Germany, Netherlands, Spain. A National Eurocoin (NE) is obtained by projecting the bandpassed GDP on a set of regressors, which are the linear combinations of national variables contained in the Thomson Financial Datastream used by the Bank of Italy (e.g. we will project German gross domestic product on 39 German variables contained in the Dataset). We look at the Eurocoin approach also as an indirect method to aggregate national data several months ahead of the official euro-area GDP estimate, with the aim to determine medium to long-run component of the European growth rate.

Firstly, we want to assess if economic time series of different countries belonging to Euro Area co-move at the European business cycle frequencies. Therefore we describe the dynamic behaviour of the *bandpassed national series* (see equation (2.1)) included in our dataset, and we also show their comovements. To do this, the evolution of Pearson correlation coefficients between each country cycle and the Euro-zone cycle (Eurocoin indicator) and national volatility can be measured. Secondly from sub-section 2.3.1 to 2.3.4, we present our national indicators based on the Eurocoin approach by ex post estimation.

We observe, in the table 2.4 below, that Germany is the most represented country in our Dataset, while Netherlands is the least.

Table 2.4 National variables contained in Thomson Datastream

Geographic Area	***Variables***
Belgium	14
Finland	2
France	23
Germany	39
Greece	1
Italy	22
Netherlands	5
Spain	25
UK	4
USA	7
Euro Area	15
TOTAL	***157***

In table 2.5 that follow, we analyze both links of the different national bandpassed medium to long run components of the growth with respect to Euro Area GDP in terms of RMSE (table 2.5/A) and among the European countries (2.5/B) in terms of correlation. We observe that France, Germany and Italy business cycles are the more correlated to the European one: the French cycle well approximates the European also in terms of RMSE in 1995-2010 (e.g. the RMSE between badnpassed French and Euro Area GDP is equal to 0.18).

Table 2.5/A – National versus Euro Area bandpassed components

	RMSE
Italy	0.31
Belgium	0.27
France	0.18
Germany	0.28
Netherlands	0.31
Spain	0.40

Table 2.5/B - Comovements in European countries in terms of correlation of the bandpassed MLRG

	Euro Area	**Belgium**	**France**	**Germany**	**Italy**	**Netherlands**	**Spain**
Euro Area	**1**	**0.89**	**0.95**	**0.94**	**0.93**	**0.90**	**0.89**
Belgium	**0.89**	**1**	**0.83**	**0.79**	**0.87**	**0.82**	**0.75**
France	**0.95**	**0.83**	**1**	**0.84**	**0.89**	**0.82**	**0.85**
Germany	**0.94**	**0.79**	**0.84**	**1**	**0.82**	**0.86**	**0.75**
Italy	**0.93**	**0.87**	**0.89**	**0.82**	**1**	**0.78**	**0.80**
Netherlands	**0.90**	**0.82**	**0.82**	**0.86**	**0.78**	**1**	**0.83**
Spain	**0.89**	**0.75**	**0.85**	**0.75**	**0.80**	**0.83**	**1**

Tables 2.5 A and B show that the French business cycle is the most similar to Euro area aggregate, the Spanish is the least similar. Performance concerning volatility is analyzed from 2003 to 2008 and from 2003 to 2009, separately, because in 2008-2010 we observe a strong recession and an high variation in volatility concerning GDP. Considering co-movements among business cycles, we observe, in terms of correlation, quite a high synchronization for Italy and France.

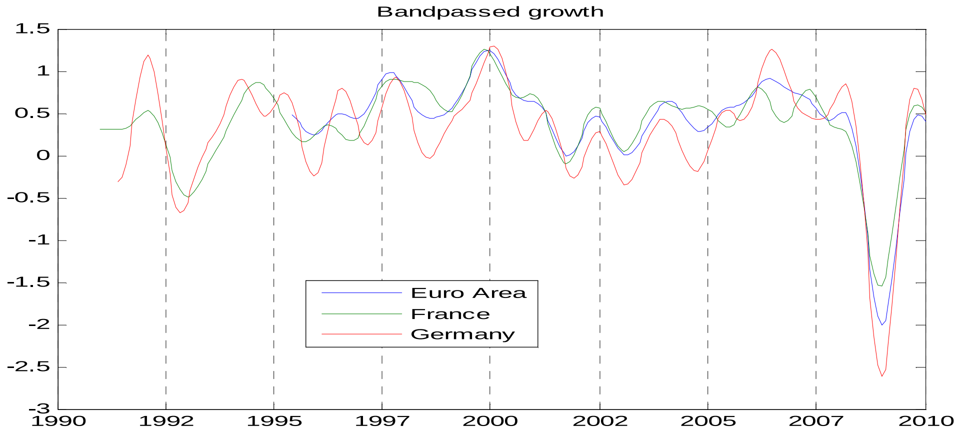
Table 2.5/C - Volatility in National Bandpassed GDP

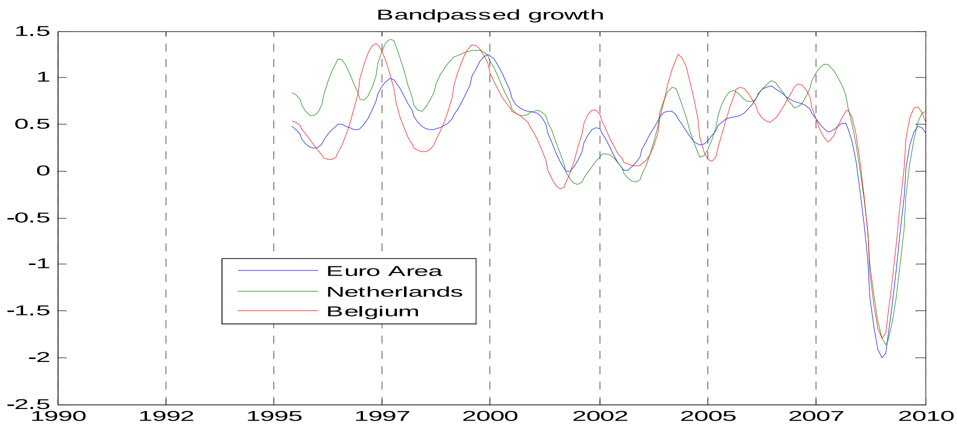
	Variance (1995,6,2007,12)	**Variance (2008,1,2010,9)**
Italy	0.14	0.87
Belgium	0.15	0.79
France	0.08	0.50
Germany	0.16	1.49
Netherlands	0.17	0.84
Spain	0.05	0.39
Euro Area	0.08	0.79

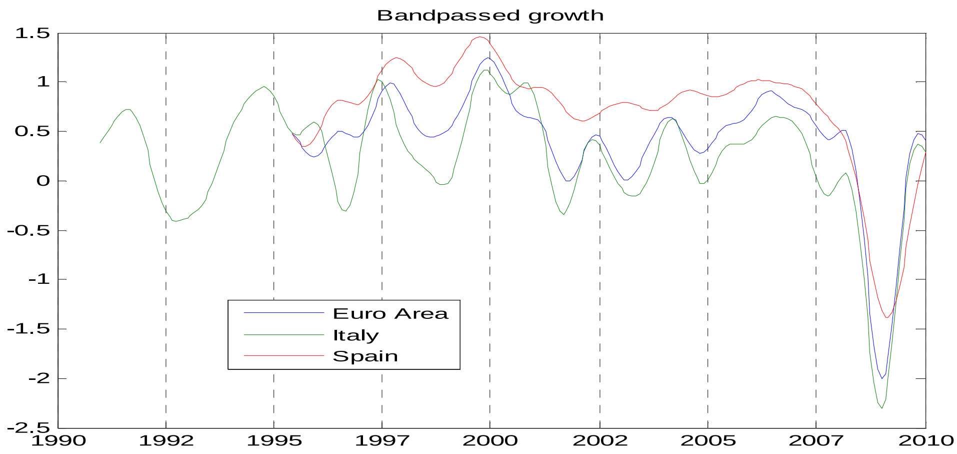
We observe the highest volatility for Germany, Italy, Netherlands, Belgium, Germany, the lower for France and Spain.

In figure 2.9 that follows, different national bandpassed growth rates are compared.

Figures 2.9 – Comparing National bandpassed growth rates







2.3.1 Why a National Eurocoin and Ex-post Results

Why build a national Eurocoin? A National Eurocoin is built because a national flash estimate of GDP is released by Eurostat about six weeks after the end of the reference quarter. A national indicator can produce real-time monthly estimate of underlying quarterly growth trends. Today, the Eurocoin is only used to outline aggregate European growth. We describe and estimate a *national Eurocoin* by projecting national and Euro Area GDP on European and national factors, which are linear combinations of the variables contained in the Thomson Financial Datastream used by the Bank of Italy.

We focus on medium to long-run component of the growth (MLRG), i.e. the smoothed component of the GDP growth rate obtained by removing the fluctuations of a period shorter than or equal to one year, and it bears no relationship to any definition of trend.

By applying the Eurocoin approach to national series, we develop some strategies to obtain smoothing of a stationary time series so as to avoid the occurrence of end-of-sample deterioration, producing monthly estimate of GDP growth, purified from erratic components (short-run fluctuations). In fact, it is a well-known result in literature that band-pass filter could deteriorate at the end of the sample. In this section, we will compare ex post data (that are generated for the same dataset that was used to estimate the parameters of the model), obtained by using our national indicators, to bandpassed GDP.

Our goal is achieved by using the generalized dynamic factor model (today used in calculating Euro MLRG by Eurocoin index, see chapter 1), for the following countries: Belgium, Italy, France, Germany, Netherlands, Spain. We decide to analyze different countries that have a different number of national variables in the dataset. We have a minimum of 5 variables for Netherlands and a maximum of 39 for Germany.

Disaggregated models with in sample estimates that we build in this paragraph will be tested in chapter 3, par.3.3, by pseudo real time simulations. In synthesis, *in sample (ex post)* estimates are generated for the same dataset that was used to estimate the parameters of the model. Otherwise, in *real time evaluation* we examine model accuracy: estimates are obtained simulating the situation one would have faced at the end of each month in terms of data availability. The latter sample will be used to build pseudo real time estimates.

We can obtain three different type of national indicator, by the strategies (methods) indicated below:

1) Projection of national GDP on European factors (so that we can use all the 157 variables contained in the dataset and used by Bank of Italy to build Eurocoin);

2) projection of Euro Area GDP on factors built by national variables (e.g. projection of European GDP on factors built by the 23 variables included in the Thomson Financial Datastream, see table 2.4);

3) projection of national GDP on common national factors (e.g. projection of French GDP on French common factors)".

Ex post estimate is looked at in this section by analyzing the in-sample 1995-2002.In chapter 3, the 2003-2010 period will be analyzed in real time at the end of the sample. *The period 1995-2010* is investigated in this thesis to analyze business cycles. Then, we will consider *n* geographic area: in table 2.6 below we have n=11, national common factors and smoothed growth being respectively identified as:

$F^{(n)}_t$ and $c_t^{(n)}$ for each area considered in our analysis.

At the end of the sample, the value of $\hat{c}_t^N$ (the national medium to long run component of GDP), with the coefficients A_i and the factors F_t, is so estimated:

$$\hat{c}_t^N = A_1 F_{1T} + A_2 F_{2T} + \dots A_m F_{mT}$$

In this section we will compare our national ex post indicators with filtered growth by country (our empirical target) that is built using a bandpass filter (see equation (2.1)). These filters work well in the middle of the sample, but they perform badly at the beginning and at the end of the sample, since it is worthwhile to develop the disaggregated indicators that we analyze more deeply in sub-section 2.3.2-2.3.5. In the following tables we outline the main descriptive statistics in 1995-2010 for bandpassed growths. We apply the final estimates of the growth rate as reported in GDP "vintage" available in February 2011. The same holds true for all other monthly series, as vintages we use for most of the monthly macroeconomic variables in the Thomson Financial Datastream are not available.

Ex post estimate is looked at in this section by analyzing the in-sample 1995-2002 for every country. In chapter 3, the period 2003-2010 will be analyzed in real time with the end of the sample. In table 2.6 below we compare in sample data for every country, obtained by applying the Eurocoin approach to national series, to bandpassed GDP. The RMSE is calculated following the equation (2.2). In table 2.7 we compare ex post national estimates with Eurocoin indicator (that regards the whole Euro Area aggregate)

Table 2.6 – Correlation and RMSFE among Ex post Indicators and National bandpassed GDP in 1995-2002

	France	Germany	Italy	Spain	Belgium	Netherlands
Projection of national GDP on European factors	**0.81**	**0.62**	**0.69**	**0.84**	**0.90**	**0.82**
RMSFE	**(0.19)**	**(0.30)**	**(0.30)**	**(0.15)**	**(0.18)**	**(0.24)**
Projection of national GDP on national factors	**0.83**	**0.71**	**0.88**	**0.89**	**0.69**	
RMSFE	**(0.18)**	**(0.27)**	**(0.20)**	**(0.13)**	**(0.30)**	**----**
Projection of Euro Area GDP on national factors	**0.82**	**0.64**	**0.78**	**0.75**	**0.57**	**----**
RMSFE	**(0.19)**	**(0.34)**	**(0.30)**	**(0.41)**	**(0.35)**	

Table 2.7 - National indicators: Correlation with Eurocoin in the period 1995-2010

	France	Germany	Italy	Spain	Belgium	Netherlands
Projection of national GDP on European factors	**0.99**	**0.96**	**0.98**	**0.89**	**0.98**	**0.98**
Projection of Euro Area GDP on national factors	**0.94**	**0.84**	**0.77**	**0.89**	**0.72**	**----**
Projection of national GDP on national Factors	**0.94**	**0.80**	**0.72**	**0.85**	**0.67**	**-----**

Concerning *Ex post* results (tables 2.6 and 2.7 above) and the figures that follow in next sub-sections obtained by using generalized dynamic factor model for national indicators, we observe that:

- the projection of GDP on national factors seems useful only for countries with more than 10 variables in the Thomson Financial Datastream (that is used to calculate Eurocoin); using less than 20 variables, we are not able to eliminate erratic components in business cycle.

- when national variables contained in the Datastream are less than 10, we are able to calculate country indicators only by projecting national GDP on the 157 European factors (see the Belgium and Netherlands cases);

- when we project a national GDP on European factors, we generally obtain an indicator similar to Eurocoin (see table 2.7). On the contrary, if we change common factors (building national factors by national variables) an indicator less correlated to Eurocoin is obtained (See table 3). This is less evident for Spain and France, where the national bandpassed GDP are respectively the least and the most correlated to the one regarding Euro Area (see table 2.5/B).

- If there is no doubt regarding building national estimates of the medium- to long-run components of GDP, it is generally better, in terms of nowcasting performance (RMSE and correlation), to project national series. It is not clear which factors use (European or national). We will investigate with regard to this point in chapter 3, section 3.3.

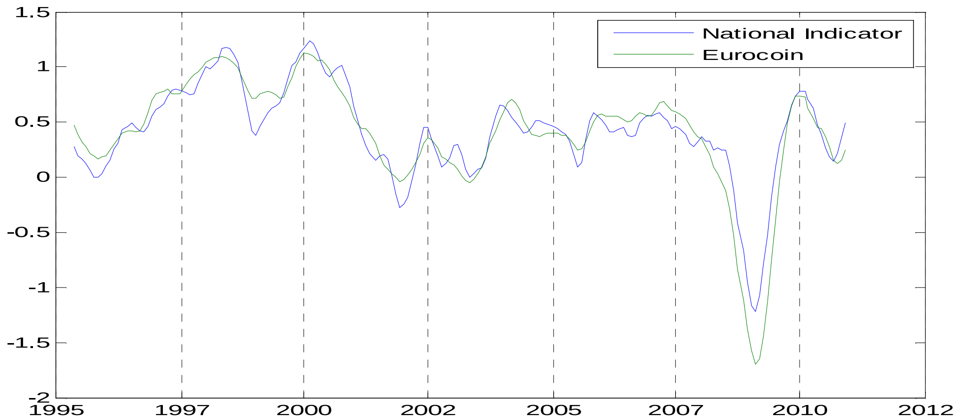
The figures that follow (from 2.10 to 2.19) show the outlook for economic activities, using Eurocoin methodology by **ex post estimates.** In addition to rigorous statistical analysis developed in sub-sections 2.3.1 for the in-sample period 1995-2002, these figures can be useful in the graphic analysis of relations among different national smoothed growths with a large period of time .

In chapter 3, we will test the strategies most indicated to outline the “country index” that we have also named “National Eurocoin”.

2.3.2 The French and German Cases

We name the indicator that we build using French data FRANCOIN, following the three different steps above indicated. It is an estimate based on past and future values of GDP. On the contrary, bandpass filter is a centered and symmetric moving average, and since it is less reliable for most recent data, it produces end of sample problems (deterioration and truncation). For values up to time t within the sample, say for 12 periods away from time T (the end of the sample) , the band-passed series, obtained by setting the missing values equal to the average of y_t (the rate of GDP growth), is a reliable estimate of c_t (the MLRG), with insignificant revisions as new data arrive. On the contrary, at the end of the sample the estimate of c_t may be misleading[27].

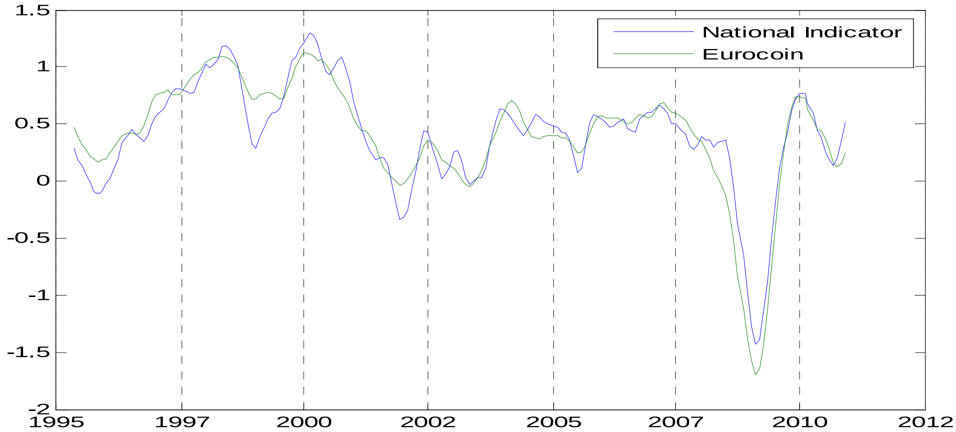
Figure 2.10 Projection of French GDP on national factors: the French case



While in figure 2.10 we have projected bandpassed French growth on the national common factors, linear combination of the 23 French variables available in the Thomson Financial Datastream (TFD), in figure 2.11 that follows, we project Euro Area GDP on the same French factors: the correlation between the relative MLRG and Eurocoin is equal to 0.94 (See table 2.7).

[27] See *New Eurocoin, A tutorial Note* (Bank of Italy)[27].

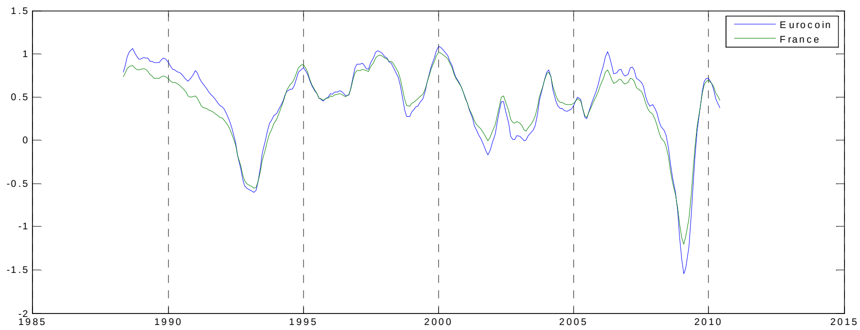
Figure 2.11 Projection of Euro GDP on national factors: the French case



In figure 2.12 we project French GDP on European common factors, linear combination of the 157 variables in the TFD, and the correlation with Eurocoin is equal to 0.99.

A national indicator thus calculated is closely correlated to Eurocoin in terms of RMSFE and correlation with our target in the period 1995-2002; from the results in table 2.6 above, it seems best strategy to use.

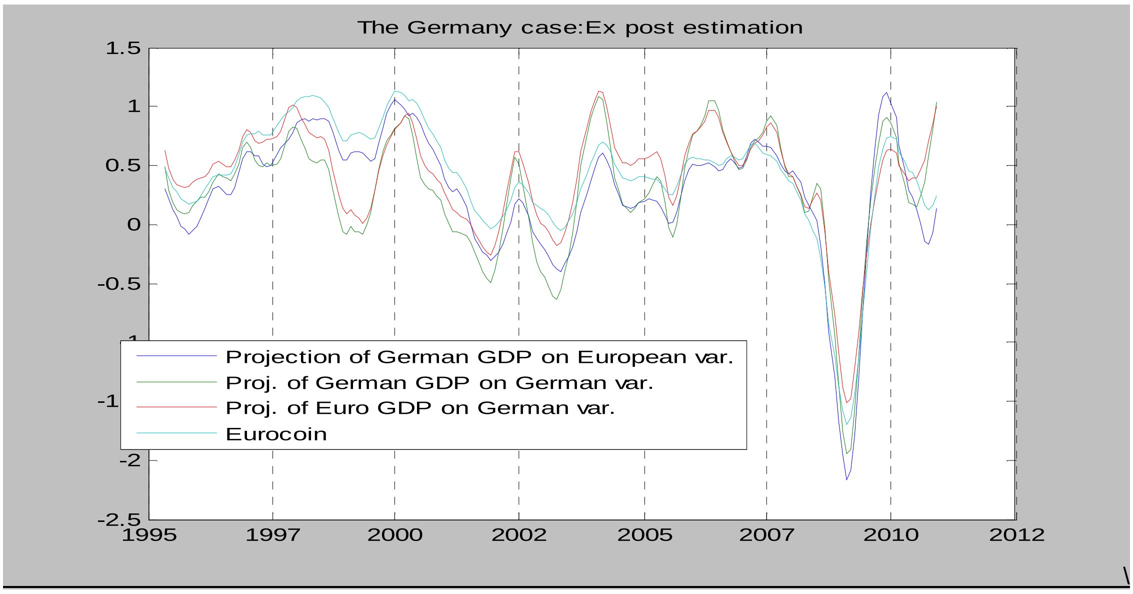
Figure 2.12: Projections of National GDP on European factors: the French case



In synthesis, the role of common factors in filtering data seems very relevant. In fact, both projecting Euro GDP and French data on French factors, we have a national indicator that is quite different from Eurocoin. On the contrary, if we project French data on European data, we have an indicator very similar to Eurocoin.

In figure 2.13 we show German growth obtained by the projection of bandpassed Germany GDP on German factors (linear combination of the 39 variables contained in the datastream, see table 2.4).

Figure 2.13 : An example: German MLRG



Also for Germany, the projection of the National GDP on European common factors produces an indicator very similar to Eurocoin.

2.3.3 The Belgium and Netherlands Cases

In this section, we use Belgium data to obtain three different types of national indicators, using generalized dynamic factor model (as Eurocoin index). Through graphic analysis, we will assess the strategies that are more likely to outline a MLRG by country:

1) Projection of national GDP on European factors (so we use all the 157 variables contained in the dataset and used by Bank of Italy to build Eurocoin);

2) projection of Euro Area GDP on Belgium factors built by 14 national variables;

3) projection of Belgium GDP on common national factors.

Figure 2.14

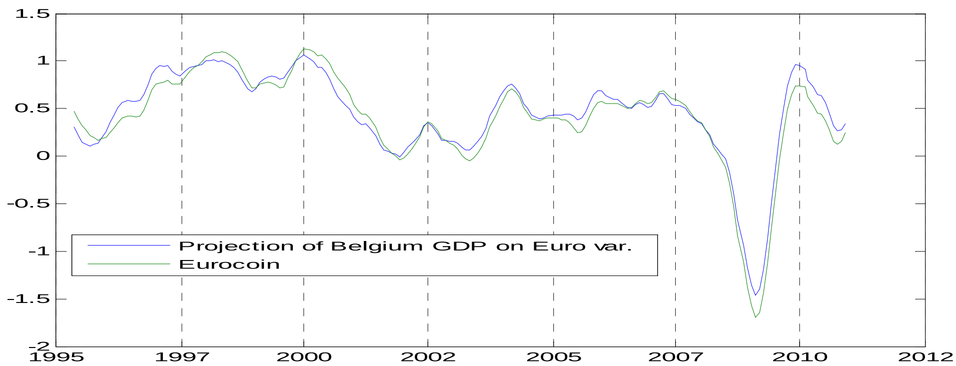


Figure 2.15

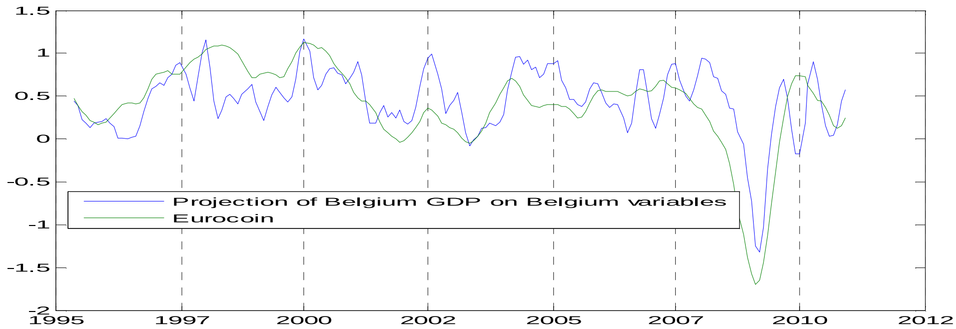
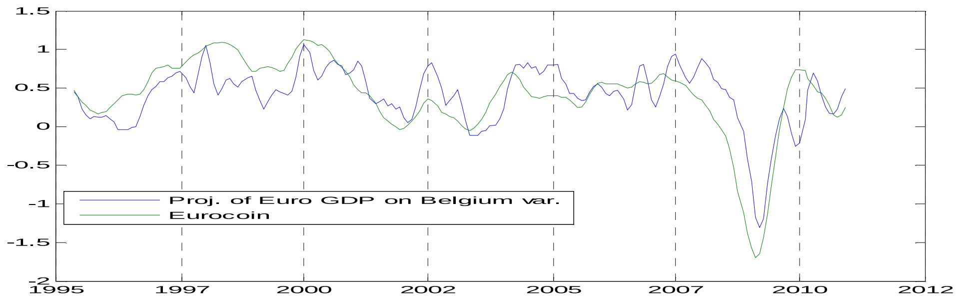


Figure 2.16



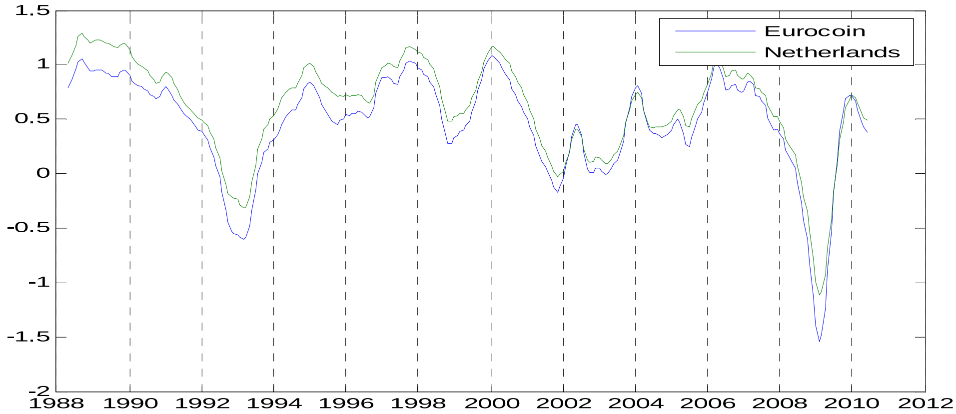
Considering figures from 2.14 to 2.16, we observe that only in figure 2.14 we obtain an indicator smoothed and free from short run fluctuations, local and sector-specific shocks and from errors in the measurement of GDP, and highly reliable at the end of the sample; it probably depends on the fact that for Belgium, only 14 variables are available in our Datastream (Thomson Financial dataset), and they are not sufficient to build smoothed factors.

Therefore, the most useful choice to outline a Belgium index seems to project Belgium GDP on Euro Area data: we can thus build our common factors by using all the 157 variables contained in our Datastream.

A Belgium Indicator based on the Eurocoin approach can therefore be obtained by the projection of the Belgium bandpassed GDP on a set of regressors (smoothing factors), which are the linear combination of the 157 European variables contained in the Thomson Financial Datastream and used by the Bank of Italy.

Since only 5 variables are available for the Netherlands, and these are not sufficient to build some smoothing factors, we will build a national index projecting the Netherlands' bandpassed GDP on Euro Area common factors: the correlation with Eurocoin is equal to 0.98 (see table 2.7) as the two indicators are very similar.

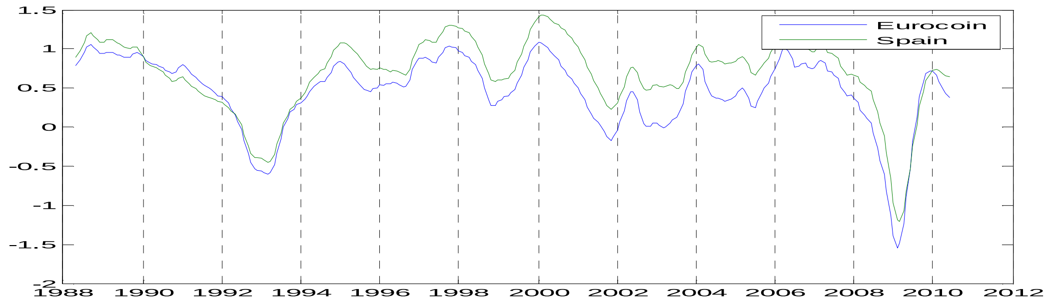
Figure 2.17: A Netherlands national indicator of busines cycle



2.3.4 Spanish and Italian Cases

In this section we produce ex-post estimation for Spanish and Italian medium to long run growth rates (MLRG), we follow to use Eurocoin approach and the generalized dynamic factor model.

Figure 2.18 Projecting Spain's GDP on European factors



In figure 2.18 we observe a really smoothed Spanish growth, while we don't observe the same positive performance by adopting the strategies considered in 2.19 and 2.20.

Figure 2.19 Projecting Spain's GDP on Spanish factors

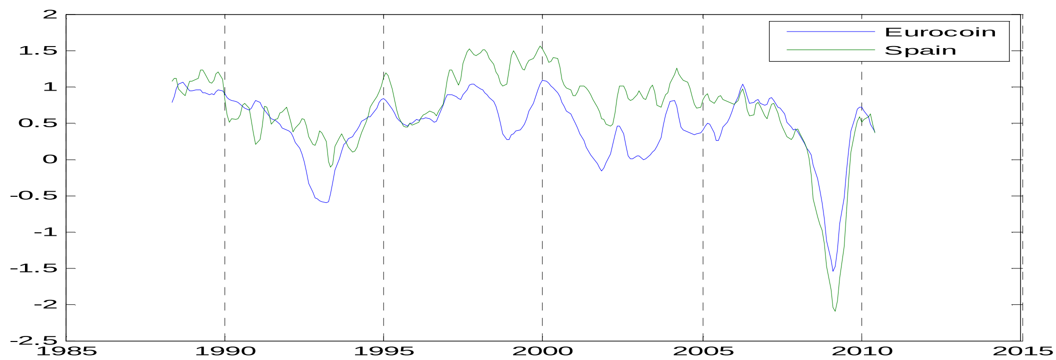
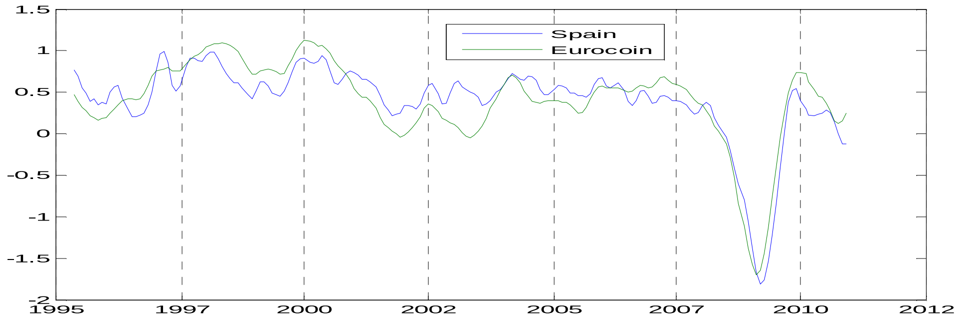


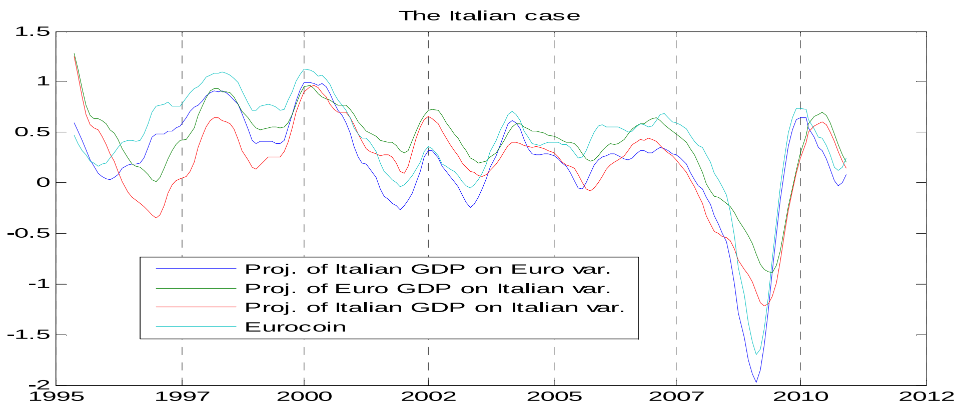
Figure 2.20 Projecting Euro Area GDP on Spanish factors



Both projecting Euro Area and Spain GDP on Spanish factors, we have a national indicator that is quite different from Eurocoin. By the in-sample (ex post) analysis, we observe that the projection of Spanish GDP on European factors (figure 2.18) seems particularly useful to eliminate erratic components, contrary to the strategy adopted in 2.20. In chapter 3, analyzing Spanish pseudo real time evaluations, we will assess if these hypothesis can be confirmed.

In the next table, it follows the Italian case, to estimate the ex-post national MLRG, by following the three different strategies tested in this section for the other countries.

Figure 2.21



Also for Italy, we observe that the three indicators developed are quite smoothed. By using European common factors, the Italian MLRG seems more similar to Eurocoin indicator (table 2.7).

A synthetic conclusion of the exercise can be that the role of common factors in filtering and estimating data is relevant at least as that concerning GDP data to be projected (in chapter 4 we will explain in detail the main results outlined in this research).

2.4 Combining Real and Financial Variables

In this section we show a tentative improving Eurocoin methodology, dividing the 157 variables contained in Thomson Financial Datastream in real and financial variables (econometric techniques to combine forecasts are presented in chapter 1).

We asses (by the methodology that we explain in par. 2.4.1) the following data groups containing "*real economic activity variables*":

- Surveys;
- Leading Indicators;
- Demand Indicators;
- Industrial production;
- Foreign Trade (Import, Export);
- Employment Indexes;
- Wages Indicators;
- Producer price Indexes.

The following data groups will be considered as containing "*financial variables*":

- Exchange rates;
- Money Supply;
- Spreads;
- Benchmark bond;
- Standard and Poor's Index.

This subdivisions among real economy and financial economy is substantially confirmed in Forni et al. (2003). There is a large literature in Macroeconomics suggesting that financial variables are good predictors of inflation and real economic activity, but empirical evidence is mixed. For a review of the empirical literature, see Stock and Watson (2001). This is clearly a puzzle for economic and finance theory. In this section and in 3.4, we assess how financial and real variables help in forecasting medium to long run components of the European growth.

We present a combination (using regression method to determine the relative weights) of real and financial MLRG that we name "*Combining Eurocoin*", useful also to analyze the impact of real and financial data in estimating smoothed GDP.

Ex post estimate is looked at in this section by analyzing the in-sample 1995-2002. In chapter 3, the period 2003-2010 will be analyzed in real time with the end of the sample.

Figure 2.22 MLRG obtained projecting GDP respectively on real and financial variables

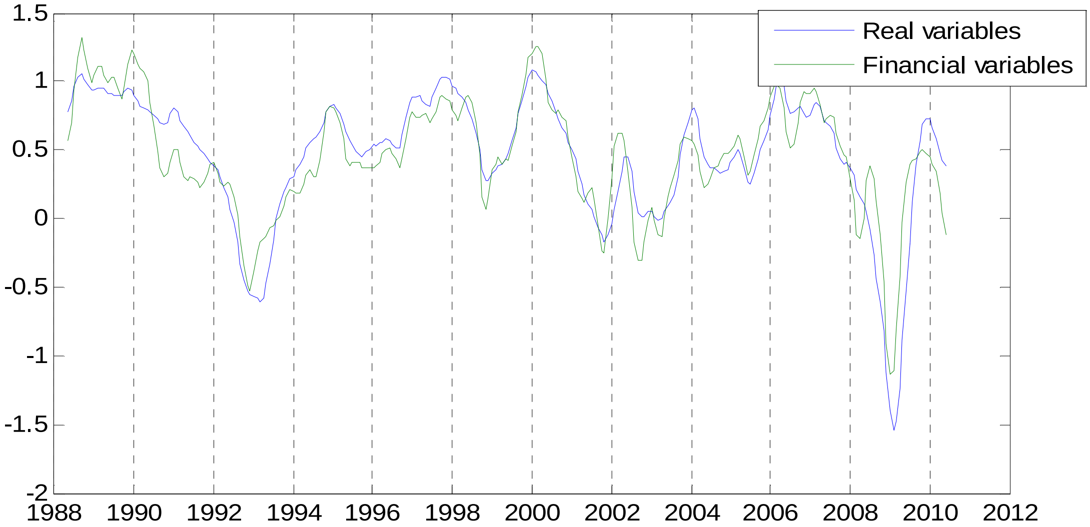
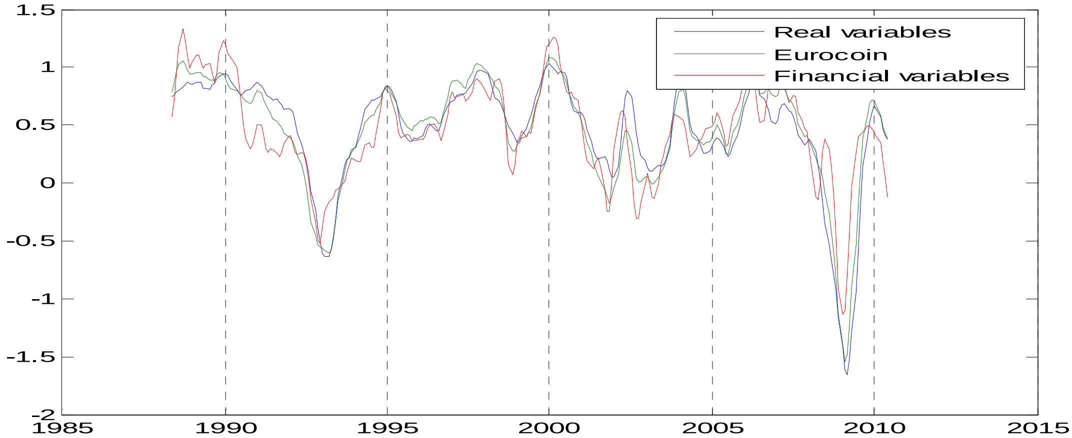


Figure 2.23 MLRG built with real and financial variables versus Eurocoin



2.4.1 The Econometric Methodology to Combine Real and Financial Variables

In a classic dynamic factor model (par 1.3, cap.1), considering the scalar time series variable Y_t to forecast and let X_t be the N-dimensional time series of candidate predictors, it is assumed that (X_t, Y_{t+h}) admits a factor model with r common latent factors F_t:

$$X_t = \Lambda F_t + \varepsilon_t$$

$$Y_{t+h} = \beta_F^{'} F_t + \beta_\omega^{'} \omega_t + \varepsilon_{t+h} \qquad (2.3)$$

where ε_t is an $N \times 1$ vector of idiosyncratic disturbances, h is the forecast horizon, ω_t is an $m \times 1$ vector of observed variables (i.e. lags of Y_t) useful, with F_t, to forecast Y_{t+h}. As we show in chapter 1, the value of c_t, with the coefficients A_i, at the end of the sample is so estimated:

$$\hat{c}_T = A_1 F_{1T} + A_2 F_{2T} + \ldots . A_m F_{mT}$$

In this research, we use two groups of common factors on which the GDP is projected: F_i and R_i (i = 1 ... m) will be respectively the common factors relevant to the prediction of "real MLRG" and "financial MLRG" (figure 2.22, 2.23), obtained by projecting Euro Area GDP respectively on real and financial variables. *Monthly weights* to combine the two smoothed growth indicators will be obtained *in real time* by the regression method (see par. 3.4 – chapter 3).

The methodology that we develop in this section can be so summarized as in:

$$X_t^F = \Lambda F_t + \varepsilon_t^F \tag{2.4}$$

$$X_t^R = \Lambda R_t + \varepsilon_t^R \tag{2.5}$$

$$Y_{t+h}^F = \beta_B' F_t + \beta_\omega' \omega_t^F + \varepsilon_{t+h}^F \tag{2.6}$$

$$Y_{t+h}^R = \beta_R' R_t + \beta_\omega' \omega_t^R + \varepsilon_{t+h}^R \tag{2.7}$$

The medium to long-run growth (that we name "Combining Eurocoin") will be equal to:

$$\hat{c}_T = \alpha_0 + \alpha_1 (A_{1F} F_{1t} + A_{2F} F_{2t} + \ldots . + A_{mF} F_{mT}) + \alpha_2 (A_{1R} R_{1T} + A_{2R} R_{2T} + \ldots . A_{mR} R_{mT})$$

Or, considering the lags as in (2.3), we have:

$$\hat{c}_T = \alpha_0 + \alpha_1 Y_{t+h}^F + \alpha_2 Y_{t+h}^R \,.$$

The comparison between the medium to long-run components, obtained through the traditional method Eurocoin, and the combination specified above, can offer some suggestions to improve the current estimate of New Eurocoin and a more specific knowledge with regard to real and financial economic activities.

2.5 Estimating GDP components: Ex-Post Results

In this section we shortly develop medium to long run growth rate (MLRG) of GDP components in Euro Area.

It seems useful to produce a monthly estimation of the smoothed GDP components (consumption, investment, exports, imports), as these national flash estimates are released by Eurostat about ten weeks after the end of the reference quarter.

The innovation of this research are some procedures, based on the Eurocoin approach and generalized dynamic factor model, that are tested in real time in section 3.5. The aim is to estimate MLRG for GDP components (consumption, investment, exports, imports) with regard to Euro Area aggregate and some specific national case. Results concerning European aggregate show that it is not always clear if it is more useful to apply national or European variables to project Expenditure Components by dynamic factor model: it is proposed to test different strategies.

These estimates for Euro Area will also be obtained in this section implementing Eurocoin approach, therefore projecting the relative bandpassed series on European latent factors.

The main technical reason why it is worthwhile to develop the disaggregated indicators that we present ex post in this section is the following:

the alternative approaches to eliminate the high frequency noise seasonal and short-run (see Baxter and King, 1999, Christiano and Fitzgerald, 2003) perform well in the middle of the sample, but they work badly at the beginning and end of the sample at the end of the sample.

The analysis that we carry out determines the contribution of different *medium to long run components of GDP* in the European business cycle and our results in real time will be carefully compared to data filtered by band-pass approach (see chapter 3, par 3.2.1). It is worthwhile to analyze if Household Consumption, Investment, Imports and exports have a weak or strong volatility, and we use Eurocoin methodology to calculate these components (today Eurocoin is only used to calculate GDP cycle).

Ex post estimate is looked at in this section by analyzing the in-sample 1995-2002. In chapter 3, the period 2003-2010 will be analyzed in real time with the end of the sample.

<u>The strategy used</u>

Projection of Euro Area GDP components (Consumptions, Investment, Foreign Trade) on European factors.

Variables that we project are:

1) Final Consumption Expenditure;
2) Household Consumptions;
3) Government Expenditures;
4) Gross Capital Formation;
5) Exports of goods and services;
6) Imports of goods and services

The aggregate Final Consumptions is composed by 2) households Consumptions and 3) Government Expenditure.

Some results

In the tables that follow, based on sample data, we show the structure of correlations and volatility for macroeconomic variables that influence the classical equation of gross domestic product in National Accounting: GDP = *Cons + Invest + Var. Stocks + Export – Import*

These variables are built using generalized dynamic factor model and these models will be tested in comparison with bandpassed GDP in chapter 3.

Table 2.8 - Correlation among bandpassed Expenditure Components

	GDP	Final Consumptions	Household Consumptions	Investments	Export	Import
GDP	-----	0.74	0.78	0.63	0.93	0.89
Final Consumptions	0.74	-------	0.98	0.44	0.53	0.58
Household Consumptions	0.78	0.98	------	0.48	0.60	0.66
Investments	0.63	0.44	0.48	------	0.49	0.79
Export	0.93	0.53	0.60	0.49	------	0.95
Import	0.89	0.58	0.66	0.79	0.95	-----

Table 2.9 - Bandpassed Expenditure Components

	MEAN (1995,6,2007,12)	MEAN (2008,1,2010,9)
GDP	**0.55**	**- 0.22**
Final Consumptions	**0.49**	**0.09**
Household Consumptions	**0.50**	**- 0.04**
Investments	**0.71**	**----**
Exports	**1.50**	**- 0.34**
Imports	**1.50**	**- 0.15**

Table 2.10 - Volatility in Bandpassed GDP components

MLRG	Variances (1995,6,2007,12)	Variances (2008,1,2010,9)
GDP	0.08	0.79
Final Consumptions	0.03	0.02
Household Consumptions	0.06	0.07
Investments	0.89	------
Exports[28]	1.30	11.7
Imports[29]	1.09	9.44

Euro Area medium to long run growth rate appears highly correlated with that concerning smoothed Export growth rate and relatively linked to Household Consumptions: the second ones have lower volatility among the different GDP components. In the figures from 2.24 to 2.27 we outline (Ex post) Euro Area GDP components growth by the Eurocoin approach and generalized dynamic factor model. In chapter 3 we will test the goodness of these models (par. 3.5) in real time.

Figure 2.24 - Consumptions

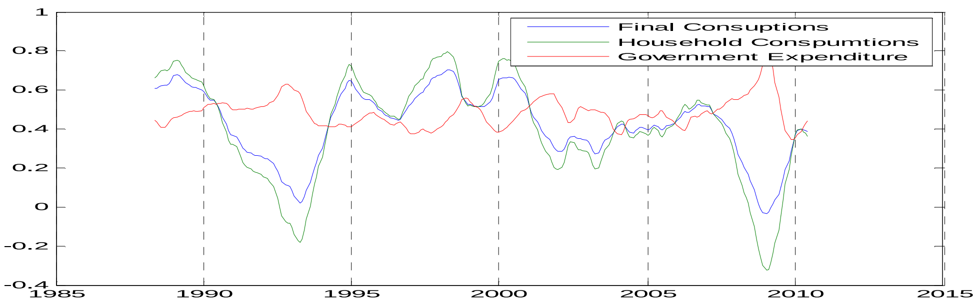
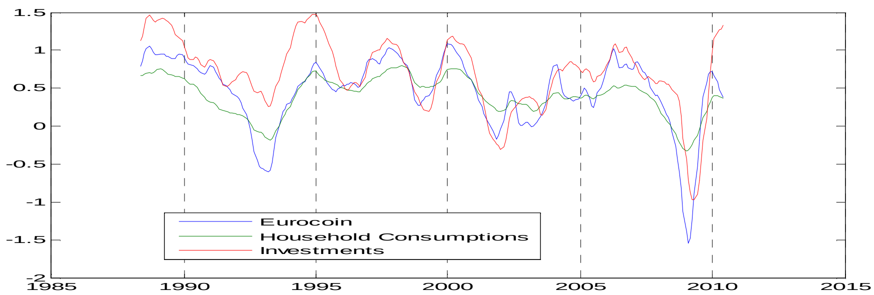


Figure 2.25 - Households Cons. and Investments versus Eurocoin



[28] Until 2008 Q3 the variance was equal to 2.06: the crisis in 2008-2010 strongly increases Export volatility.

[29] Until 2008 Q3 Import variance was equal to 1.7: the crisis in 2008-2010 strongly increases Import volatility.

Figure 2.26 - Foreign Trade MLRG

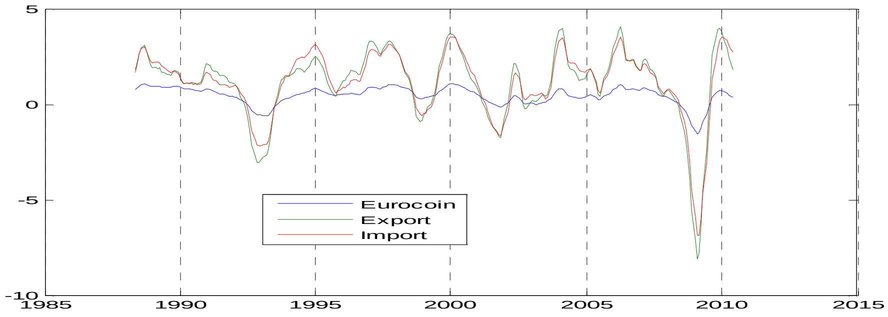
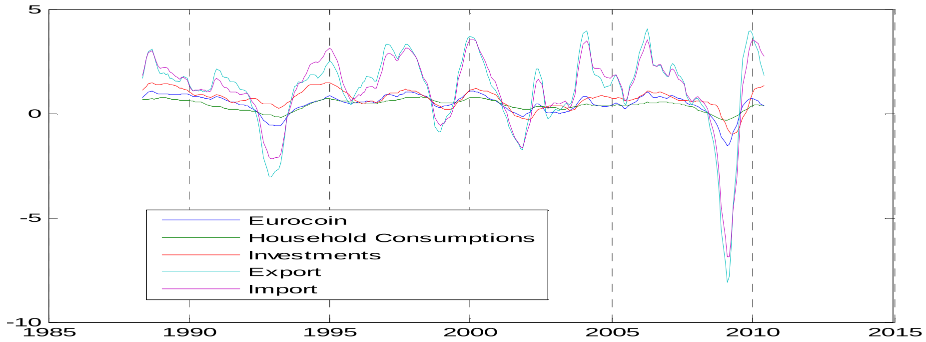


Figure 2.27 - MLRG in GDP components



Annex 3 - The Variables Used in Building Regressors

In *chapter 2* we propose a new theoretical framework for disaggregated business cycle analysis based on Generalized dynamic factor model. We project the bandpassed GDP (or GDP components) on a set of regressors, which are the linear combinations of the variables contained in the *Thomson Financial Datastream*.

The variables are available with different delays: around the 20-th of month T, when the indicator for month T is produce. Surveys, daily series and Financial Variables are generally available up to time T , thus with no delay, Industrial Orders and Car Registrations up to $T - 1$ and Industrial Production indicators up to $T - 2$ or $T - 3$.

This end-of-sample imbalance problem can be solved by one of these methods:

- forward realignment, that is simply shifting forward the variables that are available with delay;
- in alternative, the variables that are not available can be nowcasted by using ARMA models or the EM (expectation-maximization)[30] algorithm.

The data source that we use are the following:

- Surveys:

They are conducted each month by European countries and EC (European Commission) and concern economic climate, stock of finished goods, manufacturing, retail and construction sectors as well as consumers.

- OECD Composite Leading Indicators (Trend Restored)

OECD CLIs are aggregate time series which show a leading relationship; aggregating the components series into the CLI reduces the risk of "false signals", changes in the indicator due to irregular and erratic movements.

The composite leading indicator (CLI) is an aggregate time series showing a consistent leading relationship with the growth cycles of key macro-economic indicators in a country. CLI is constructed by aggregating component series selected according to multiple criteria, such as: economic significance, cyclical correspondence and data quality. These component series cover a wide range of short-term indicators such as observations or opinions about economic activity, housing permits, financial and monetary data.

The OECD approach concerning Composite Leading Indicators (CLIs) is designed to provide early signals of turning points in business cycles (peaks and troughs), and provide qualitative rather than quantitative information (forecasts) on short-term economic movements.

The CLIs can be presented in various forms. The *trend restored CLI* is composed of the trend of the reference series (in index form or natural units) and the amplitude adjusted CLI. The CLI is adjusted to ensure that its cyclical amplitude on average agrees with that of detrended reference series, facilitating analyses of business cycle.

[30] Dempster et al. (1977)

- IIP (Index of industrial production)

It is a typical measure of real activity. The index measures monthly evolution of the volume of industrial production (in Italy it excludes Construction sector, in Germany and Belgium it is calculated including Construction), together with their breakdown by final destination (Consumption, Investment and Intermediate goods).

- Demand indicators

This data source includes:

1) retail sales;
2) household consumptions (manufactured and durable goods);
3) new passenger car registrations;
4) wholesale trade turnover.

- Wages

This source includes the following variables:

1) Unit Labour Costs (IMF source);
2) Hourly Rate in Industry (OECD source).

- Data concerning Unemployment rate and job vacancies

- Data concerning foreign trade (Import - Export)

- Producer Price Index

These indices measure the average change in the price of goods and services as they leave the place of production valued at basic prices. The precise way PPIs are constructed depends on who and for what they are meant to be used. In this context, national practices may differ and these differences impact on international comparability between countries especially on aspects such as:

- the weighing and aggregation systems;

- the sampling selection;

- the treatment of product substitution (to take account of the price behaviour of new and disappearing products that generally do not have the same price);

- the treatment of seasonal items as supplies and prices of some products are subject to strong seasonal variations.

- Exchange rates among Euro, $, £

- Monetary aggregates

They are compiled by central banks on the basis of surveys of monetary and financial institutions: they measure the amount of money circulating in an economy, and, the national contribution to the Euro Area.

- Spread

In finance, a **credit spread** is the difference in yield between different securities, due to different credit qualities. It reflects the additional net yield an investor can earn from a security with more credit risk relative to one with less credit risk. Credit spread of a particular security is often quoted in relation to the yield on a credit risk-free benchmark security or reference rate. This is a variable only considered, in our dataset, for UK, USA, Italy, Germany

- Interest rates

- Benchmark indicators of stock market and concerning securities

Annex 4 – List of Monthly Macroeconomic Indicators by Country

Applying, in this thesis, a dynamic factor model to the construction of European coincident indicators of the business cycle by economic sector and by expenditure components of GDP (e.g Household Consumptions) require that economic time series of different countries and sectors strongly co-move at business cycle frequencies.
In this annex all the indicators used in building disaggregated MLRG in this thesis are specified by country and by source.

Business and consumer confidence indicators

Germany

CONSUMER SURVEY: MAJOR PURCH. OVER NEXT 12 MONTHS	EUROPEAN COMMISSION
MFG. EXCL. F.B.T.: ORDERS ON HAND - BAL.	IFO INSTITUT
BUSINESS CLIMATE INDEX	IFO INSTITUT
ASSESSMENT OF BUSINESS SITUATION	IFO INSTITUT
BUSINESS EXPECTATIONS	IFO INSTITUT
BUSINESS CLIMATE INDEX: MANUFACTURING	IFO INSTITUT
ASSESSMENT OF BUSINESS SITUATION: MANUFACTURING	IFO INSTITUT
BUSINESS EXPECTATIONS: MANUFACTURING	IFO INSTITUT
ASSESSMENT OF BUSINESS SITUATION: CONSTRUCTION	IFO INSTITUT
ASSESS. OF BUS. SITUATION: RET. TRADE (INCL. MOT.VEH., PETR.ST.)	IFO INSTITUT

Belgium

BUSINESS INDICATOR SURVEY - ECONOMY	BANQUE NATIONALE DE BELGIQUE
BNB BUS. SVY. - TRADE - NOT SMOOTHED	BANQUE NATIONALE DE BELGIQUE

Spain

ECONOMIC SENTIMENT INDICATOR - SPAIN	EUROPEAN COMMISSION
INDUSTRY SURVEY: ORDER BOOK POSITION - SPAIN	EUROPEAN COMMISSION
CONSUMER SURVEY: MAJOR PURCH. OVER NEXT 12 MONTHS - SPAIN	EUROPEAN COMMISSION

France

SURVEY: MANUFACTURING - SYNTHETIC BUSINESS INDICATOR	I.N.S.E.E.
SURVEY: INDUSTRY - RECENT OUTPUT TREND	I.N.S.E.E.
SURVEY: INDUSTRY - ORDER BOOK & DEMAND	I.N.S.E.E.

SURVEY: MANUFACTURING OUTPUT LEVEL - GENERAL OUTLOOK	I.N.S.E.E.
SURVEY: INDUSTRY – PROBABLE OUTPUT TREND	I.N.S.E.E.

Italy

ISAE CONSUMER SURVEY: ECONOMIC CLIMATE INDEX - GENERAL	ISAE
ISAE BUSINESS SVY.: STOCKS OF FINISHED GOODS	ISAE
ISAE BUSINESS SVY.: PRODUCTION IN NEXT 3MOS.	ISAE
ISAE CONSUMER SURVEY: ECONOMIC CLIMATE INDEX - FUTURE	ISAE

Netherlands

CBS CONSUMER CONFIDENCE SURVEY: INDEX	CENTRAAL BUREAU VOOR DE STATISTIEK

Business and consumer confidence indicators (seasonally adjusted)

Germany

NEW REGISTRATIONS - CARS	KRAFTFAHRT-BUNDESAMT
RETAIL SALES EXCLUDING CARS	STATISTISCHES BUNDESAMT, WIESBADEN
WHOLESALE TRADE TURNOVER, NOMINAL	STATISTISCHES BUNDESAMT, WIESBADEN

Belgium

NEW PASSENGER CAR REGISTRATIONS	A.C.E.A.
RETAIL SALES	INSTITUT NATIONAL DE STATISTIQUE

Spain

REGISTRATIONS: PASSENGER CAR	MINISTERIO DE ECONOMIA Y HACIENDA

France

NEW CAR REGISTRATIONS	I.N.S.E.E.
HOUSEHOLD CONSUMPTION - MANUFACTURED GOODS, RETAIL GOODS	I.N.S.E.E.
HOUSEHOLD CONSUMPTION - MANUFACTURED GOODS	I.N.S.E.E.
HOUSEHOLD CONSUMPTION - DURABLE GOODS	I.N.S.E.E.

Italy

NEW PASSENGER CAR REGISTRATIONS	ANFIA
RETAIL SALES	ISTITUTO NAZIONALE DI STATISTICA

Industrial production indices

Germany

INDUSTRIAL PRODUCTION INCLUDING CONSTRUCTION	DEUTSCHE BUNDESBANK
INDUSTRIAL PRODUCTION - INTERMEDIATE GOODS	DEUTSCHE BUNDESBANK
INDUSTRIAL PRODUCTION: MANUFACTURING	STATISTISCHES BUNDESAMT
INDL.PROD.: CHEMICAL PRODUCTS	STATISTISCHES BUNDESAMT
INDL.PROD.: RUBBER AND PLASTIC PRODUCTS	STATISTISCHES BUNDESAMT
INDL.PROD.: BASIC METALS	STATISTISCHES BUNDESAMT
INDL.PROD.: MFG, COMPUTER, ELECCL. & OPT. PRDS., ELECL.EQP.	STATISTISCHES BUNDESAMT, WIESBADEN

Belgium

INDUSTRIAL PRODUCTION INCL. CONSTRUCTION	INSTITUT NATIONAL DE STATISTIQUE
INDUSTRIAL PRODUCTION - INTERMEDIATE PRODUCTS	INSTITUT NATIONAL DE STATISTIQUE
INDUSTRIAL PRODUCTION - MANUFACTURING	INSTITUT NATIONAL DE STATISTIQUE

Spain

INDUSTRIAL PRODUCTION EXCLUDING CONSTRUCTION	EUROSTAT
INDUSTRIAL PRODUCTION - INTERMEDIATE GOODS	INSTITUTO NACIONAL DE ESTADISTICA (INE)
INDUSTRIAL PRODUCTION - CAPITAL GOODS	INSTITUTO NACIONAL DE ESTADISTICA (INE)

INDUSTRIAL PRODUCTION - OTHER NON-METAL MINERAL PRODUCTS	INSTITUTO NACIONAL DE ESTADISTICA (INE)
INDUSTRIAL PRODUCTION (WDA)	MINISTERIO DE ECONOMIA Y HACIENDA
INDUSTRIAL PRODUCTION - MACHINERY & MECHANICAL EQUIPMENT	INSTITUTO NACIONAL DE ESTADISTICA (INE)

France

INDUSTRIAL PRODUCTION - ENERGY	I.N.S.E.E.
INDUSTRIAL PRODUCTION - MANUFACTURING	I.N.S.E.E.

Italy

INDUSTRIAL PRODUCTION EXCLUDING CONSTRUCTION	EUROSTAT
INDUSTRIAL PRODUCTION	ISTITUTO NAZIONALE DI STATISTICA
INDUSTRIAL PRODUCTION: CONSUMER GOODS	ISTITUTO NAZIONALE DI STATISTICA
INDUSTRIAL PRODUCTION: INVESTMENT GOODS	ISTITUTO NAZIONALE DI STATISTICA

Netherlands

INDUSTRIAL PRODUCTION EXCLUDING CONSTRUCTION	CENTRAAL BUREAU VOOR DE STATISTIEK

EMU

INDUSTRIAL PRODUCTION: MANUFACTURING	EUROSTAT
INDUSTRIAL PRODUCTION EXCLUDING CONSTRUCTION (EA16)	EUROSTAT
INDUSTRIAL PRODUCTION: MANUFACTURING - PULP, PAPER & PAPERBOARD	EUROSTAT
INDUSTRIAL PRODUCTION: MANUFACTURING - OTHER CHEMICAL PRODUCTS	EUROSTAT
INDUSTRIAL PRODUCTION: MANUFACTURING - BASIC METALS	EUROSTAT
INDUSTRIAL PRODUCTION: MANUFACTURING - GENERAL-PURPOSE MACHINERY	EUROSTAT

France

INDUSTRIAL PRODUCTION - INVESTMENT GOODS	I.N.S.E.E.
INDUSTRIAL PRODUCTION - CONSUMER DURABLE GOODS	I.N.S.E.E.
INDUSTRIAL PRODUCTION EXCLUDING CONSTRUCTION	I.N.S.E.E.

Wages and Employment indices

Germany

UNIT LABOUR COSTS, RELATIVE NORMALIZED	IMF INTERNATIONAL FINANCIAL STATISTICS
VACANCIES (PAN BD FROM JAN 1994)	DEUTSCHE BUNDESBANK

Italy

HOURLY RATES IN INDUSTRY	MAIN ECONOMIC INDICATORS,COPYRIGHT OECD

France

UNEMPLOYMENT RATE: PERSONS UNDER 25 YEARS OLD	EUROSTAT

Spain

EMPLOYMENT PROMOTION CONTRACTS: IN PRACTICE	MINISTERIO DE ECONOMIA Y HACIENDA
UNEMPLOYMENT: REGISTERED - UNDER 25 (SILE)	MINISTERIO DE ECONOMIA Y HACIENDA
JOB VACANCIES (EP) $METHODOLOGY BREAK FROM MAY 2005	INSTITUTO NACIONAL DE EMPLEO (INEM)

Producer Price Index

Germany

PPI: NON-DURABLE CONSUMER GOODS	EUROSTAT
PPI: INDUSTRIAL PRODUCTS, TOTAL, SOLD ON THE DOMESTIC MARKET	STATISTISCHES BUNDESAMT, WIESBADEN
PPI: ENERGY	EUROSTAT
PPI: INDUSTRY (EXCLUDING CONSTRUCTION)	EUROSTAT

France

PPI: INTERMEDIATE GOODS	EUROSTAT

Italy

PPI	ISTAT
PPI: NON-DURABLE CONSUMER GOODS	EUROSTAT
PPI: ENERGY	EUROSTAT

Spain

PPI - MANUFACTURING INDUSTRY	INSTITUTO NACIONAL DE ESTADISTICA (INE)
PPI - CONSUMER GOODS, DURABLES	INSTITUTO NACIONAL DE ESTADISTICA (INE)
PPI - CONSUMER GOODS, NON-DURABLES	INSTITUTO NACIONAL DE ESTADISTICA (INE)
PPI - CAPITAL GOODS	INSTITUTO NACIONAL DE ESTADISTICA (INE)
PPI - INTERMEDIATE GOODS	INSTITUTO NACIONAL DE ESTADISTICA (INE)
PPI - ENERGY	INSTITUTO NACIONAL DE ESTADISTICA (INE)

Belgium

PPI - INDUSTRY (EXCLUDING CONSTRUCTION)	BANQUE NATIONALE DE BELGIQUE
PPI - ENERGY: ELECTRICITY, GAS, STEAM, WATER	BANQUE NATIONALE DE BELGIQUE
PPI - INTERMEDIATE GOODS	INSTITUT NATIONAL DE STATISTIQUE
PPI - CONSUMER GOODS	BANQUE NATIONALE DE BELGIQUE
PPI - INVESTMENT GOODS	BANQUE NATIONALE DE BELGIQUE

Netherlands

PPI: INTERMEDIATE GOODS	EUROSTAT
PPI	CENTRAAL BUREAU VOOR DE STATISTIEK

EMU

PPI: INDUSTRY (EXCLUDING CONSTRUCTION) (EA16)	EUROSTAT
CPI - ALL ITEMS (HARMONISED) (EA16)	EUROSTAT

Finland

PPI	STATISTICS FINLAND
PPI - MANUFACTURING	STATISTICS FINLAND

Greece

PPI: DOMESTIC MARKET - MINING, QUARRYING & MANUFACTURING INDUSTR	NATIONAL STATISTICAL SERVICE - GREECE

External Transactions

IMPORTS (CIF) **NETHERLANDS**	CENTRAAL BUREAU VOOR DE STATISTIEK
IMPORTS (CIF) **SPAIN**	MAIN ECONOMIC INDICATORS,COPYRIGHT OECD
IMPORTS (CIF) **GERMANY**	DEUTSCHE BUNDESBANK
EXPORTS FOB **BELGIQUE**	BANQUE NATIONALE DE BELGIQUE
EXPORTS FOB **FRANCE**	FRENCH CUSTOMS
EXPORTS FOB **SPAIN**	MAIN ECONOMIC INDICATORS,COPYRIGHT OECD

Financial variables

MONEY SUPPLY - M3 (CONTINUOUS SERIES)	DEUTSCHE BUNDESBANK	***GERMANY***
MONEY SUPPLY - M2 (CONTINUOUS SERIES) (PAN BD FROM 1991)	DEUTSCHE BUNDESBANK	***GERMANY***
MONEY SUPPLY - M1 (NATIONAL CONTRIBUTION TO M1)	BANQUE DE FRANCE	***FRANCE***
MONEY SUPPLY - M3 (NATIONAL CONTRIBUTION TO M3)	BANQUE DE FRANCE	***FRANCE***
MONEY SUPPLY: M1 - ITALIAN CONTRIBUTION TO THE EURO AREA	BANCA D'ITALIA	***ITALY***
MONEY SUPPLY: M3 - ITALIAN CONTRIBUTION TO THE EURO AREA	BANCA D'ITALIA	***ITALY***
MONEY SUPPLY: M1 (EP)	EUROPEAN CENTRAL BANK (ECB)	***EMU***
MONEY SUPPLY: M3 (EP)	EUROPEAN CENTRAL BANK (ECB)	***EMU***
FTSE ITALIA MIB STORICO		***ITALY***
DJ EURO STOXX		***EMU***
DAX 30 PERFORMANCE		***GERMANY***
FRANCE CAC 40		***FRANCE***
S&P 500 COMPOSITE		***UNITED STATES***
FTSE 100		***UK***
IBEX 35		***SPAIN***
EURIBOR 3 MONTH		***EMU***
UK INTERBANK 3 MTH (LDN:BBA)		***UNITED***

		KINGDOM
BD BENCHMARK 3 YEAR DS GOVT. INDEX		***GERMANY***
BD BENCHMARK 10 YEAR DS GOVT. INDEX		***GERMANY***
US BENCHMARK 3 YEAR DS GOVT. INDEX		***UNITED STATES***
UK £ TO EURO (ECB)		***UK***
US $ TO EURO (ECB)		***UNITED STATES***
FRANCE BENCHMARK BOND 10 YR (DS)		***FRANCE***
SPAIN BENCHMARK BOND 10 YR (DS)		***SPAIN***
ITALY spread		***ITALY***
GERMANY spread		***GERMANY***
UNITED STATES spread		***UNITED STATES***
UK spread		***UK***
REAL EFFECTIVE EXCHANGE RATE - CPI BASED	MAIN ECONOMIC INDICATORS, OECD	***GERMANY***

Chapter 3: Pseudo Real Time Estimation

3.1 Introduction

As pointed out in chapter 1 and 2, New Eurocoin indicator provides a summary index of the medium to long-run component (MLRG) of the GDP only for the whole Euro area aggregate. The main contribution of this research is to propose some procedures to estimate the following European GDP components:

- Sectoral smoothed growths;
- MLRG concerning some countries belonging to the Euro Area;
- Expenditure components.

This section focuses on the pseudo real time estimating performance of these disaggregated indicators. Period by period, real time estimates are reported here; estimates have been obtained simulating the situation one would have faced at the end of each month in terms of data availability. It is worthwhile to specify that the term *"pseudo"* indicates that the matrix of data we use does not consider the revision to the series. The same series are cropped over and again to take account of restriction to the set of information. Experiments conducted in this chapter use 6 generalized principal components. This is a "pseudo-realtime execise" since we use the "final" data in the estimation and not the first releases that are not available for all data series.

As it is explained on the **€-coin** website, in most cases (daily and monthly data) revisions are non-existent or relatively minor. For GDP (the only quarterly series used), revisions are more relevant. However, given recent change in national account practices in the euro area (chain linking and different treatment of FISIM) it is not clear what meaning one could attach to a real time exercise that encompasses a sharp change in the definitions and rules adopted by National Statistical institutes to compute GDP.

Experiments conducted in this chapter use 6 generalized principal components, the number estimated over the whole sample period [1 T]. The exercises we develop use the estimates $\hat{c}_t(t+h)$, of each disaggregated indicators at time t using the data from 1 to t + h, h = 0, 1, 2, with t running from January 2003 to December 2010.

Real Time performance is computed by using the following steps and it gives a sense of how well the model has gone at the end of the sample:

1. Select a date near the end of the sample;
2. Estimate model using data up to that date;
3. Use estimated model to produce some forecasts/estimates, by using a recursive or a rolling window, or their combination[31], R*ecursive simulation scheme* proceeds as

[31] See Clark, T. and M. McCracken (2009).

follows: the initial estimation date is fixed, but additional information is added one at a time to the estimation period. Whereas, a *rolling window* is one where the length of the in-sample period used to estimate the model is fixed, so that the start and end date increase successively by one observation. Therefore, in the rolling scheme estimation sample size remains constant. On the contrary, in the recursive scheme, sample size increases every period. While the recursive window has the advantage of using all data available at a certain point in time, the rolling window skips information (see table 3.1).

In the theoretical case of infinite data series, evaluation of the medium to long-run component can easily be done by applying band-pass filtering. In reality, band-pass filter method provides a good approximation in the middle of the sample, while approximations at its ends are very poor. It is not an appropriate approach for real-time analysis. The idea of Eurocoin Indicator approach is based on the assumption that a panel of macroeconomic variables capture some information about future GDP dynamics, to perform equally well within and at the end of the sample.

In this thesis, using a large dataset of macroeconomic variables, a few smooth unobservable factors will be constructed by Eurocoin approach to describe Euro Area economy. The aims will be to estimate: sectoral MLRG; National Business Cycle Coincidence Indicators; smoothed GDP components (i.e. Household Consumptions). Each real time indicator will be compared to the target that is a band pass bilateral filter on growth rate. Target value, which is not available at the end-of-sample time T, is available with good accuracy only at time T +h, for a suitable h. As a consequence, disaggregated indicators produced at time T will be compared with the target at T produced at time T + h.

Analysis of real time performance, in this chapter will regard:

- ability of indicators to signal the correct sign of the change concerning the target. It will be measured as the difference between our indicator at time t and the approximate target at t that is obtained using data up to T, by calculating the RMSFE (root mean squared forecast error). Real time error include both uncertainty concerning future values of error term and that arising due to the fact that regression coefficients are estimated;
- ability of real time indicators to signal the correct sign of target change;
- size of the revision errors after one month;
- ability to perform well in signalling turning points in the target.

It is also important to distinguish the concept of *Out of Sample Estimation,* explicitly stating that no information outside those enclosed in your sample should be used to estimate the holdout sample from that of *Real time Performance (used in this chapter).* The latter simulates to dispose data up to a certain date so as to estimate the indicator at that time.

This chapter is organized as follows:

- in par. 3.2 we outline and test in real time the “Sectoral Eurocoin (SE);
- in par 3.3 national indicators for the smoothed GDP in Euro Area are examined;

- 3.4 shows in real time the impact of real and financial variables on estimate the medium to long run component of the growth.
- In 3.5 we assess real time performance of the smoothed growth rate for each component of GDP (Consumptions; Foreign Trade; Investments).

3.1.1 The Real Time Problem

When new data differ from previous estimates produced by the indicators, they can be revised over the estimation period. This can cause indicators to produce some false signals in real time. To evaluate how well they perform in real time, it is useful to conduct a simulated real-time forecasting experiment.

Within each month contemporaneous values of key variables are unavailable. At an arbitrary point in each month ν, data available is the information set Ω_{ν}^{n} that includes the most recent data for n monthly macroeconomic series.

The econometrician's task will be to project GDP growth $y_{\nu+h}$ for $h = 0,...,H$ based on the information set available at ν:

$$\hat{y}_{\nu+h} = Proj\left[GDP \mid \Omega_{\upsilon}^{n}\right], h = 0,..., H$$

In fact, for the euro area, a flash estimate of GDP is released by Eurostat about six weeks after the end of the reference quarter, and a full set of indicators for the second quarter of the year is not available any earlier than the flash estimate of GDP.

We can therefore consider the Eurocoin approach as a method to forecast smoothed GDP, since Eurocoin index produces preliminary estimates of the medium to long run component of the growth and it does not deteriorate at the end of the sample. On the contrary, flash estimates of GDP are produced some weeks after Eurocoin, while bandpass filter deteriorates at the end of the sample (so, their estimates can't be timely as per those concerning the Eurocoin approach) .

As outlined in chapter 1 and 2, New Eurocoin and the disaggregated (national, sectoral, financial and real) models built in chapter 2 of this Thesis, tested real time in this section, are based on generalized dynamic factor models. These models project c_t (the bandpassed data concerning GDP) on a set of regressors, which are linear combinations of the variables contained in the *Thomson Financial Datastream* used by the Bank of Italy. It is a well-known result in literature that band-pass filter could deteriorate at the end of the sample. New Eurocoin is a method to obtain the smoothing of a stationary time series so as to avoid the occurrence of end-of-sample deterioration, producing real-time monthly estimate of GDP growth, purified from erratic components (short-run fluctuations).

The variables contained in the Datastream (see annexes 3-4) considered in our research are used in building simulated forecasts. A *recursive window* will be used to validate our model in real time.

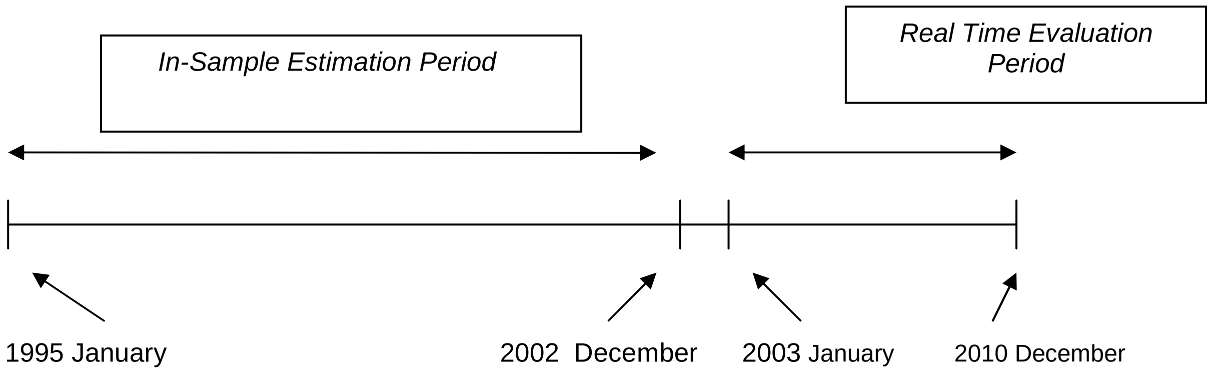


Figure 3.1 Use of In Sample and an Real Time Estimation in our research

Table 3.1 – Rolling versus Recursive window

Objective: to produce 1-,2-,3-steps-ahead forecasts for:	***Data to estimate parameters of the model in 2003***	
	Rolling window	***Recursive window***
2003 M1,M2,M3	1989M1-2002M12	1989M1-2002M12
2003 M2,M3,M4	1989M2-2003M1	1989M1-2003M1
2003 M3,M4,M5	1989M3-2003M2	1989M1-2003M2
2003 M4,M5,M6	1989M4-2003M3	1989M1-2003M3
2003 M5,M6,M7	1989M5-2003M4	1989M1-2003M4
2003 M6,M7,M8	1989M6-2003M5	1989M1-2003M5
2003 M7,M8,M9	1989M7-2003M6	1989M1-2003M6
2003 M8,M9,M10	1989M8-2003M7	1989M1-2003M7
2003M9,M10,M11	1989M9-2003M8	1989M1-2003M8
2003M10,M11,M12	1989M10-2003M9	1989M1-2003M9

Main conclusions of the research with regard to disaggregated indicators included in the real time analysis are presented in chapter 4.

3.2 Sectoral Indicators: Real Time Evaluation

In chapter 2, by implementing Eurocoin methodology, we have outlined a disaggregated analysis of the medium to long run components of GDP (MLRG) in the European economy, studying interrelations and characteristics of sectoral smoothed growth rates. The main macro sectors analyzed are: Manufacturing, Energy, Construction, Financial, Trade-Transport-Communication.

In this paragraph, we will test our sectoral models by using real time estimation.

Sectoral estimates are thus used in this section to calculate errors in real time:

1. We choose a number of observations, P, for which we produce P monthly real time estimates for the period January 2003 - December 2010; Let s = T – P;
2. Estimate the model using the dataset for t= 1,…,s;
3. Compute the estimate for the first month beyond our sample, s+1; we call this $\hat{Y}_{s+1|s}$;
4. Calculate the error, $\hat{s}_{s+1} = Y_{s+1} - \hat{Y}_{s+1|s}$ and RMSFE that quantifies uncertainty. Y_{s+1} is the monthly bandpassed sectoral estimate.
5. We will repeat steps 2-4 for the remaining months, s = T-P+1 to T-1 (November 2010).

Step 4 is changed calculating turning points and revision errors (see sub-section 3.2.2).

Among these sectoral indicators, we compare the better performances in real time with the different versions of bandpass filters (Baxter-king, Christiano-Fitzgerald, see 3.2.5).

3.2.1 Medium to Long-Run Growth of Euro Area Sectoral GDP: Band-Pass Filter Approach

Following Altissimo et al. (2006), medium to long-run growth of economic activity can be calculated by removing any fluctuations from real GDP quarterly growth, of a period shorter than or equal to one year.

Sectoral MLRG is defined considering the spectral decomposition of y_t^s, real GDP quarterly growth for the main economic sectors in the aggregate Euro Area. Assuming stationarity y_t^s can be represented as the sum of sine and cosine waves with different weights, with $\beta(L)$ that is a low-pass filter excluding short waves with a frequency equal to or higher than $\frac{\pi}{6}$, that corresponds to a period of 1 year. Using the bandpass filters (see e.g. Baxter and King, 1999, and Christiano and Fitzgerald, 2003), our medium to long run components c_t^s , for each sector

s, is the following symmetric, infinite, two sided linear combination of the sectoral GDP growth series:

$$c_t^s = \beta(L)y_t^s = \sum_{k=-\infty}^{\infty} \beta_k y_{t-k}^s, \qquad \beta_k = \begin{cases} \frac{(k\pi/6)}{k\pi} \; for \;\; k \neq 0 \\ 1/6 \quad for \; k = 0 \end{cases} \tag{3.1}$$

Then, $y_t{}^s$ has the following decomposition:

$$y_t^s = c_t^s + s_t^s \;\; \beta(L)y_t^s + [1-\beta(L)]y_t^s$$

Since $\beta(1) = 1,$ the mean μ of $y_t{}^s$ is retained in c_t while the mean of the excluded part of the sectoral GDP growth is equal to zero.

It is necessary to point out that our ideal target, being an infinite moving average, is unobservable. Therefore, since the data on sectoral GDP are finite, equation (3.1) cannot be applied in practice. So, a finite-sample version of the band-pass filter (equation 3.2 below) provides a good approximation to the ideal target at time t in the middle of the sample, and it performs badly at the beginning and end of the sample. Precisely, the performance of the sectoral Eurocoin at time t, with $t \;\leq T-12$, will be measured as the difference between our indicator at time t and the approximate target at t that is obtained using data up to T.

According to Altissimo et al. (2006), within a finite sample the following approximation of the target can be obtained, by augmenting $y_t{}^s$ with its sample mean $\hat{\mu}$ in both infinite directions:

$$c_t^{*s} = \beta(L)y_t^{*s}, \quad \text{where } y_t^{*s} = \begin{cases} y_t \quad if \quad 1 \leq t \leq T \\ \hat{\mu} \;\; if \quad t < 1 \; or \, t > T \end{cases} \tag{3.2}$$

Since y_t, the sectoral growth rate, is observed only quarterly, while we are interested in a monthly indicator of economic activity, we chose a simple interpolation to calculate the two missing points for each quarter, assuming that y_t is unchanging within a quarter.

This approximation of sectoral MLRG is reasonable in the middle of the sample over the period $\begin{bmatrix} 13 & T-12 \end{bmatrix}$.

Since estimation of the last data point in the approximate bandpassed target is bad (e.g. see figures 2.3 in sub-section 2.2.1, chapter 2), in next subsection some sectoral indicators to calculate MLRG are developed not only for the past but also in real time.

3.2.2 The Real-Time Performance

Below, a pseudo real-time evaluation of a "Sectoral Eurocoin" is outlined. Here "pseudo" refers to the fact that true real-time preliminary estimates of the GDP are not applied, but the final estimates as reported in GDP "vintage" available in February 2011. Data concerning sectoral breakdown of quarterly gross value added are available in Eurostat Statistics Database by themes and the European Central Bank Monthly Bulletin (Euro Area Statistics on line). The same holds true for all other monthly series, as vintages used for most of the monthly macroeconomic variables in the Thomson Financial Datastream are not available.

As we have explained in chapter 1, subsection 1.3.1, the truncated band-pass filter c_t^{*s} can deteriorate at the end of the sample. The main contribution to the literature in this section is the implementation of the New Eurocoin approach to different economic sectors. Monthly estimation in real time concerning sectoral medium to long-run growth will be obtained projecting sectoral bandpassed GDP on common smooth factors.

As we have described in chapter 2, section 2.2.1, monthly estimation in real time concerning sectoral medium to long-run growth $\hat{c}_t^{s}$ will be obtained projecting sectoral bandpassed GDP on common smooth factors, which are generalized principal components of current values of the variables in the dataset. Since only current values of the macroeconomic series are used, no end of sample deterioration occurs and we can improve the estimate of the sectoral MLRG at the end of the sample.

Following Altissimo et al. (2009), the criteria to be applied in pseudo-real time evaluation is established to analyze the ability of $\hat{c}_t^{s}$ for each sector *s* considered and to estimate (approximate) the truncated band-pass filter c_t^{*s}. Therefore, we analyze the performance of the "Sectoral Eurocoin" at time t, with $t \leq T-12$, by the difference between our indicator at time t and the approximate target at t that is obtained using data up to T.

Real time performance is analyzed from 2003 to 2008 and from 2003 to 2009, separately, because in 2008-2010 we observe a strong recession and an high variation in volatility concerning GDP.

Monthly length of the sample 2002-2010 is equal to T=108 for the period T-96 ≤ t ≤ T-13. We are interested in:

a) the ability of the real time indicator to approximate the target as measured by the root mean-square error $\sqrt{\frac{\sum_{t=T-96}^{T-13}\left[\hat{c}_t^{s}(t)-c_t^{*s}(T)\right]^2}{84}}$; the results of this statistic will always be indicated in the tables of this chapter as " *Rmse with respect to c**".

b) the size of the revision errors after one month, that are measured by the ratio $r=\sqrt{\frac{\sum_{t=T-96}^{T-1}\left[\hat{c}_t^{s}(t+1)-\hat{c}_t^{s}(t)\right]^2}{96}}$; to measure the revision the sample is extended up to T. The results of this statistic will be always be indicated in the tables of this chapter as "*Rmse revision errors*".

c) the ability of $\hat{c}_t^{s}(t)$ - $\hat{c}_{t-1}^{s}(t)$ = $\Delta\hat{c}_t^{s}(t)$ to signal the correct change of bandpassed variation $\Delta c_t^{*s}(T)$ (see Peasaran and Timmermann, 1992).

Table 3.2/A - Performance in real time: Sectoral Eurocoin

SECTORS	RMSE with respect to c*	RMSE with respect to c*
	2003,1,2008,1	**2003,1,2009,12**
Manufacturing + Energy	0.39	1.35
Manufacturing	0.44	1.47
Construction	0.63	0.96
Trade-Transport	0.18	0.58
Financial	0.27	0.41

Notes: Sample January 2002-December 2010; the sub-sample in the real time exercise is 2003-2009

Table 3.2/A above shows real time forecasting performance of real-time indicators with respect to a measure of the "trend-cycle GDP growth" obtained in the middle of the sample by a band pass bilateral filter on GDP growth.

During the 2008-2009 financial crisis we find a very high volatility in the bandpassed target, and real time performance decreased significantly in terms of RMSFE: this period is not considered for the analysis of revision errors[32]. Table 3.2/A-B shows that in terms of RMSE both Trade-Transport-Communication and Financial sector score better than Industrial sectors.

Table 3.2/B - Performance in real time: Sectoral Eurocoin

SECTORS	**RMSE revision errors** **2003,1,2008,1**
Manufacturing + Energy	0.05
Manufacturing	0.06
Construction	0.06
Trade-Transport	0.03
Financial	0.03

In this section, to assess the ability of $\Delta \hat{c}_t^s(t)$ to signal the correct change of the bandpassed variation $\Delta c_t^{*s}(T)$, we use the statistical test of Pesaran and Timmermann (1992), that is more deeply investigated in the Annex 2 (chapter 1).

In synthesis, Pesaran and Timmermann proposed a directional accuracy (DA) test of the hypothesis that there is no relationship between the direction of change predicted by a model and the observed change (i.e., in our case, the bandpassed sectoral target). Concerning our disaggregated estimates, if P is the proportion of times the sign of sectoral and smoothed growth rate (the approximate target) that is correctly predicted by the sectoral Eurocoin in real time, and P_star is the probability of the correct sign being estimated under the assumption that the predictor is independent from the predicted variable, we can shortly highlight, following Pesaran andTimmermann (1992), that

$$S\_n = \frac{(P - P\_star)}{\sqrt{Var(P) - Var(P\_star)}}$$

is approximately normal (see equation 1.37 and 1.39 in the annex 2).

The null hypothesis of the test (see table 3.2/C below) is that real time indicator (the predictor) and the (predicted) bandpassed target are independently distributed, i.e. the sectoral Eurocoin for each sector, does not have power in nowcasting the target, and they are independently

[32] This unprecedented depth of the recession that hit the euro area has also induced a reassessment of the Eurocoin estimation procedure to keep track of the events.

distributed; 95% and 99% critical values of a standard normal variate are, respectively, 1.96 and 2.576. We observe that PT two sided test is above the 99% critical value for:

- the sectoral aggregate "Manufacturing + Energy";
- the aggregate Trade-Transport-Communication (in the following it is indicated as TTC) that strongly rejects the null hypothesis.

PT is above the 95% critical value for Manufacturing. For the Construction sector we observe a bad performance in terms of correct prediction of sign, and for the Financial we detect that the significant level of the test is very low (high p-value).

Table 3.2/C Non-parametric Statistic of Pesaran - Timmermann (PT)

SECTORS	**PT**	**p-value of the PT test statistic**	**% Correct prediction of sign of bandpassed Δc^* 2003,1,2009,12**
Manufacturing + Energy	2.58	0.0103	0.64
Manufacturing	2.33	0.0201	0.63
Construction	- 2.07	0.0381	0.38
Trade-Transport	2.83	0.0046	0.65
Financial	-0.13	0.8971	0.49

Figure 3.3 below shows sectoral estimates in real time that we develop in this chapter; in 3.4 they are compared to the bandpassed targets.

Figure 3.3

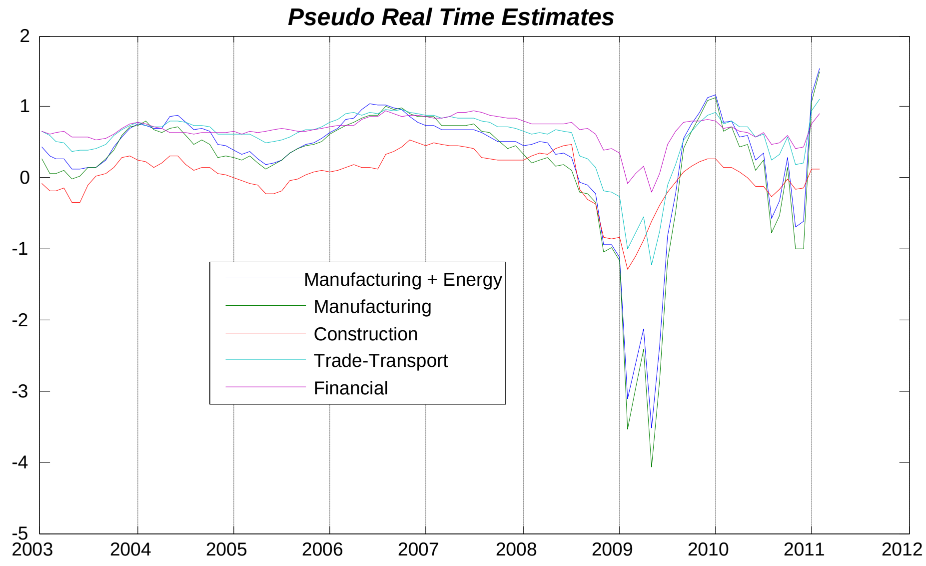


Figure 3.4/A - C_t^* and the Real Time Indicator: the Manufacturing case

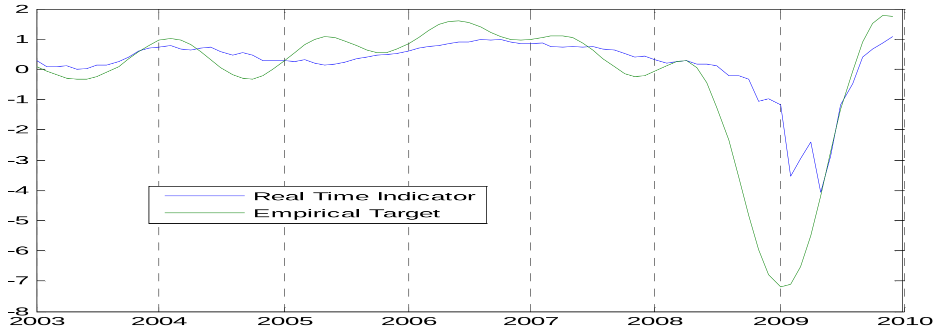


Figure 3.4/B - C_t^* and the Real Time Indicator: the Manufacturing and Energy case

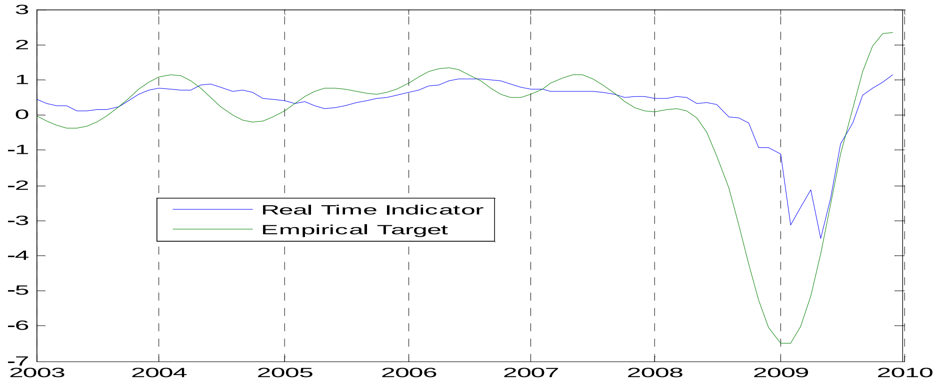


Figure 3.4/C - C_t^* and the Real Time Indicator: the Construction case

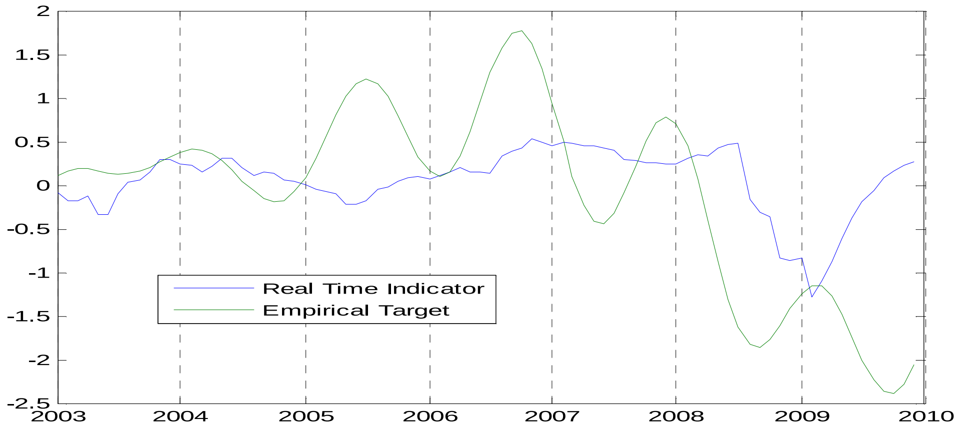


Figure 3.4/D - C_t^* and the Real Time Indicator: the Trade-Transport-Communication case

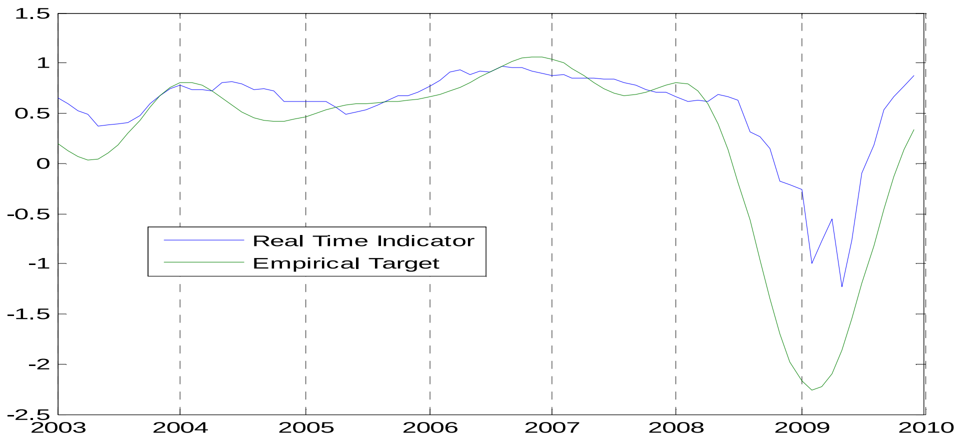
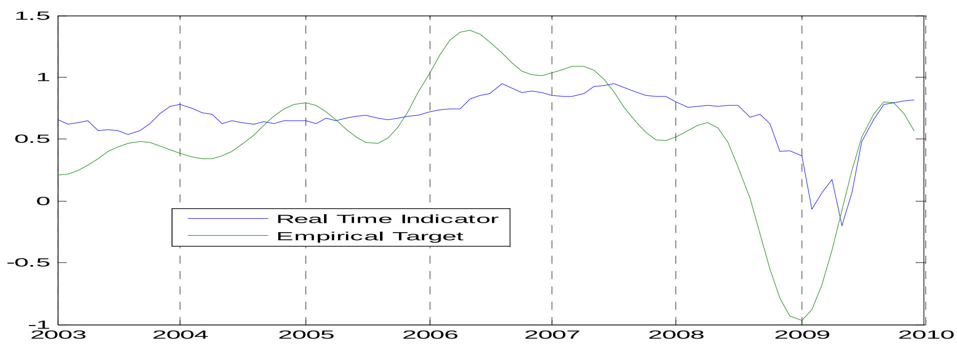


Figure 3.4/E - C_t^* and the Real Time Indicator: the Financial case



By the figures above we find that:

- for Construction and Financial, the ability of each real time indicator to estimate its target is low;
- on the contrary, we observe that the other indicators (with regard to Manufacturing, Energy, Trade, Transport, Communication) work quite well.

Main conclusions:

By sectoral real time estimation we observe, particularly, that:

- regarding the nowcast of each sectoral target C_t^* (table 3.2/A), Trade-Transport-Communication (TTC) aggregate sector scores remarkably better than Manufacturing and Construction;
- also in terms of size of revision errors after one month (table 3.2/B), TTC performs better than other sectors;

- with regard to the ability of real time indicator to track target (table 3.2/C), we observe that the TTC indicator performs fairly well. Also for the aggregate Manufacturing and Energy, PT statistic is above 2.57, 99% critical value for a two sided test. This test shows a bad performance for the Construction and Financial sectors.

3.2.3 Behaviour around Turning Points

A characteristic of the sectoral indicators that we test in this chapter, is the ability to give a correct signal of MLRG turning points in real time.
In the *Glossary for OECD Composite Predicting Indicators* it is outlined that " A turning point (TP) occurs in a series when the deviation-from-trend series reaches a local maximum (Peak) or a local minimum (Trough). Growth cycle peaks (end of expansion) occur when activity is furthest above its trend level. Growth cycle troughs (end of contraction/recession) occur when activity is furthest below its trend level. In addition, turning points should respect various censoring rules. In the simplified *Bry-Boschan* procedure (1971), used in the OECD CLI system for turning point identification, these censor rules guarantee the alternation of peaks and troughs, while ensuring that phases last not less than 9 months and cycles last not less than 2 years".
This methodology is based on the concept which focuses on fluctuations in the absolute level of economic activity; however, since this work is based on fluctuations in q-o-q growth rate, we say that an upturn (downturn) signal in $\hat{c}_t^s$ can be predicting or lagging true upturn, tolerating a four-month error.

Similarly to the NBER procedures summarized in the Bry and Boschan program (see annex 1), the following turning points are defined, for each sector in the subsample 2003-2010:

Table 3.3/A – Number of Turning Points in the Bandpassed Target

SECTORS	TOTAL TURNING POINTS	DOWNTURNS	UPTURNS
Manufacturing + Energy	4	2	2
Manufacturing	4	2	2
Construction	6	3	3
Trade-Transport	4	2	2
Financial	4	2	2

Table 3.3/B:Real time detection of turning points (TP)

SECTORS	TP Signals	Correct TP	Correct over signalled TP	Missed over all TP
Manufacturing + Energy	**4**	**2**	**2/4**	**2/4**
Manufacturing	**4**	**2**	**2/4**	**2/4**
Construction	**4**	**2**	**2/4**	**4/6**
Trade-Transport	**4**	**3**	**3/4**	**1/4**
Financial	**4**	**-**	**---**	**4/4**

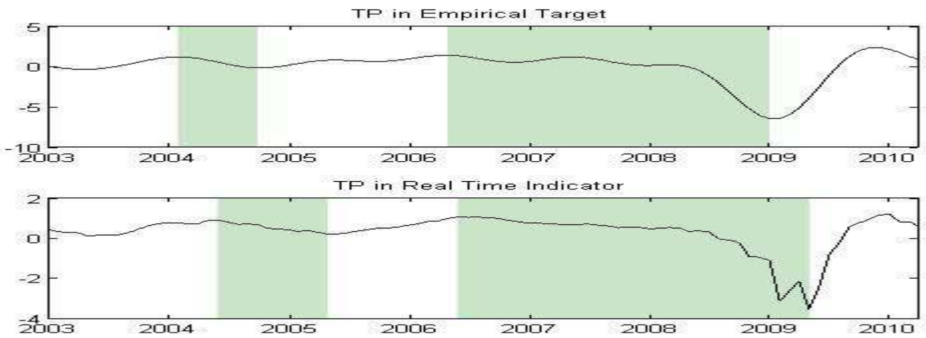
Coming to the turning points identified by each sectoral indicator, we observe that Manufacturing, Energy, Trade-Transport sectors perform fairly well in terms of missed points over all TP.

In the figures 3.5 below, we observe period characterized by recession and economic contraction in green, while the white area regards expansion phases. In each figure the Empirical target (see equation (3.2)) is compared to the real time performance.

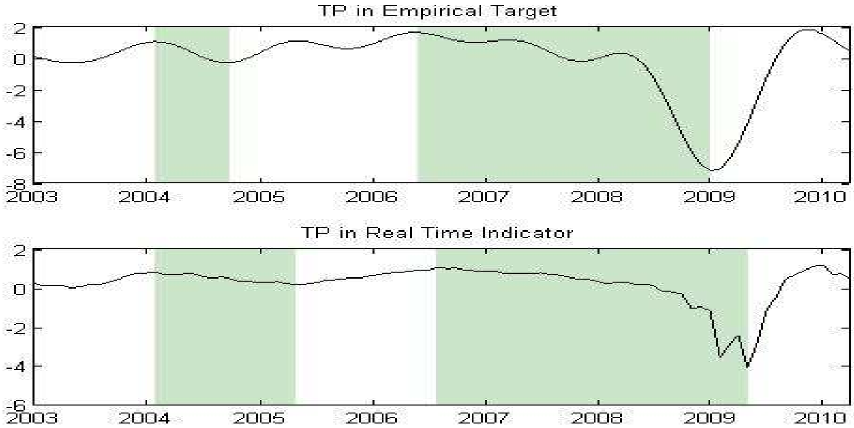
Figures 3.5: Real Time detection of Sectoral Turning points

Shade Recession/Contraction Area

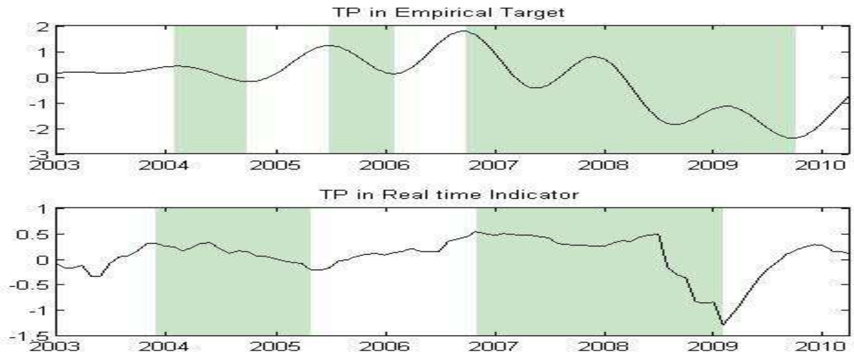
A) Manufacturing and Energy case



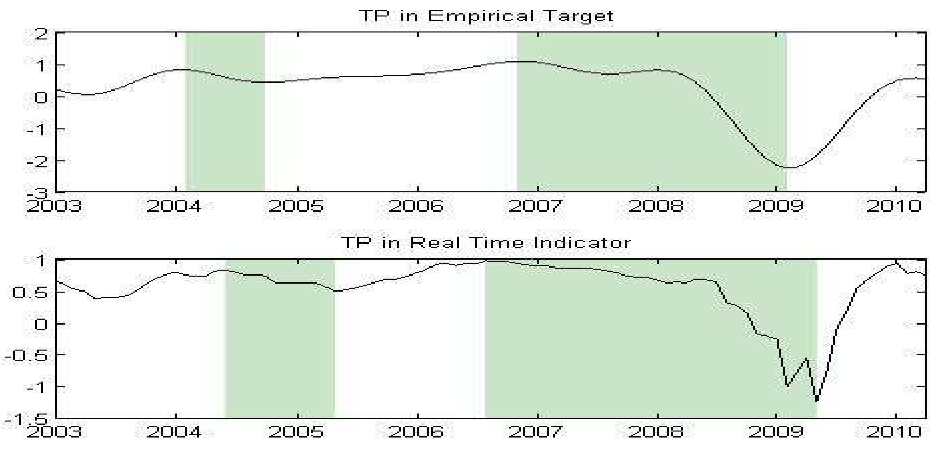
B) Manufacturing



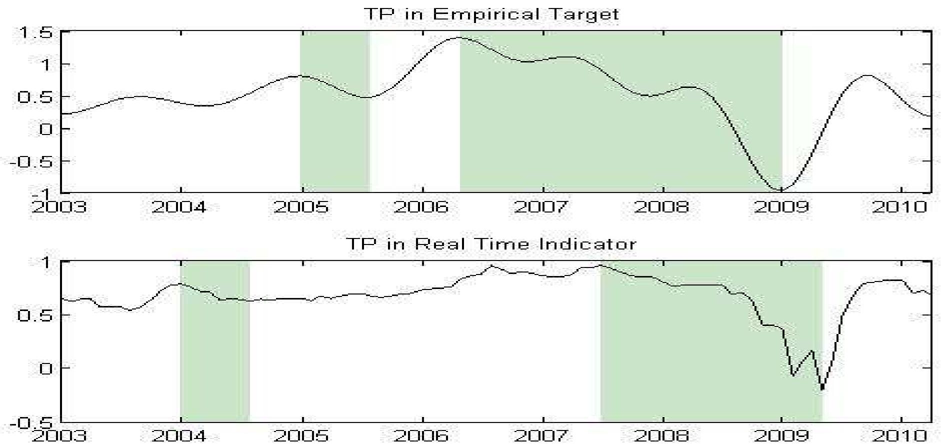
C) Construction



D) Trade-Transport-Communication



E) Financial



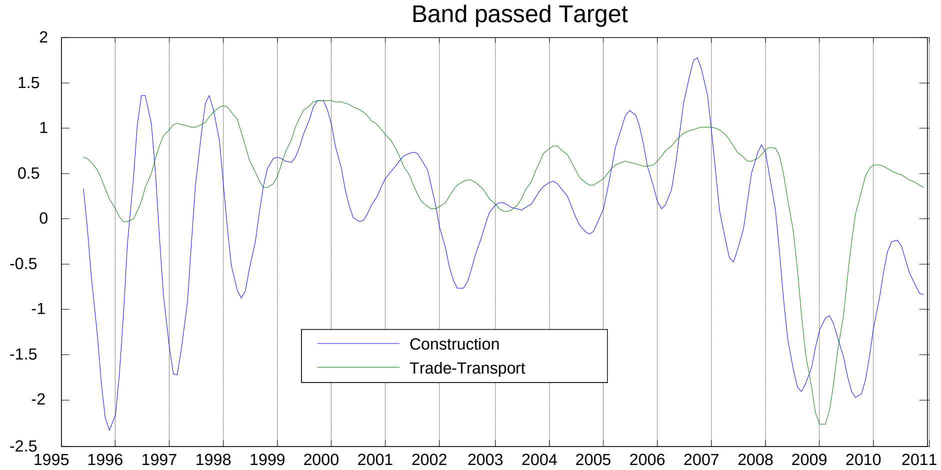
Concerning the cycle phases outlined above, we observe, in particular, the capacity of Trade-Transport Real Time Indicator to predict the recession phase in 2006 in relation to Empirical Target, while Financial indicator perform poorly in signalling cycle phases.

<u>Some final remarks</u>

Why is the Construction indicator in real time based on the dynamic factor model is not a good predictor for the relative sectoral MLRG? We identify in Construction sector a high number of fluctuations (considering the period 1989-2010) with regard to the Euro area aggregate and a strong volatility.

In figure 3.6 below we compare Construction and TTC MLRG in terms of bandpassed target.

Figure 3.6: Construction versus Trade-Transport Empirical Target



As we show below in table 3.4, volatility concerning the Construction sector, differently from the one regarding TTC, is particularly high also in the period between January 2003 and December 2007[33]. Observing Trade and Constructions pseudo real time performance (tables 3.2-3.3), we can think that the lower volatility of TTC bandpassed growth rate improves performance in real time.

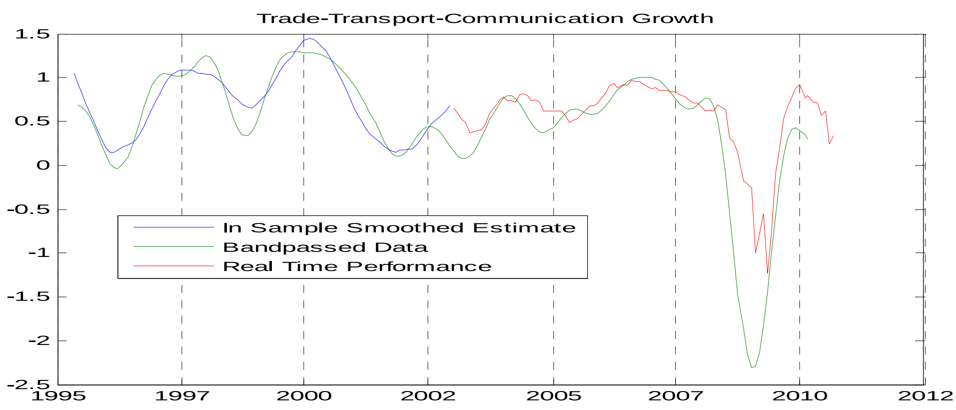
Table 3.4 - Volatility in Bandpassed Growth

SECTORS	**2003-2007**
Manufacturing	0.27
Manufacturing + Energy	0.20
Constructions	0.28
Trade-Transport-Communication	0.07
Financial	0.10

Therefore, the 157 European variables from Thomson Financial Datastream that we use in building our common factors, are particularly useful to assess a sectoral indicator for Trade-Transport-Communication, but not to predict "Construction MLRG".

In figure 3.7 below, we observe pseudo real time performance of Trade, Transport and Communication branches with ex post (in sample) and with bandpassed data, showing the capacity of the Sectoral Eurocoin to be timely and updated at the end of the sample, contrary to bandpassed data.

Figure 3.7



[33] During financial crisis, In 2008-2010, volatility is particularly high for all sectors.

Finally, it seems interesting to note that we tested also to eliminate production industrial data from our dataset, with the goal to build "Service" sector estimation. Results that we obtained show that by eliminating industrial data, to build some estimates of the smoothed growth rate for Financial, Trading, Transport and Communication sectors, doesn't change the structure of the common factors and, in terms of RMSE, we obtain some results similar to the ones showed in table 3.2/A.

3.2.4 Aggregating Sectoral Data

Monthly estimation in real time concerning sectoral medium to long-run growth is obtained by the projection of sectoral bandpassed GDP on a set of regressors which are the linear combination of the 157 variables contained in the *Thomson Financial Datastream* and used by the Bank of Italy. In this section we build our real time aggregated and sectoral estimates by projecting the following aggregated data:

- Manufacturing, Energy and Construction data: we produce an "Industry" Indicator of sectoral growth;
- Trade-Transport-Communication-Financial are aggregated by producing a "Service" MLRG (medium to long run growth rate). Aggregated indicators are defined in the following as "macro indicators". Results of our test are shown in the tables 3.5.

Table 3.5/A – RMSFE among Real Time Indicators and Sectoral Bandpassed GDP

MACRO SECTORS	**2003,1,2008,1**	**2003,1,2009,12**
Industry	**0.37**	**1.26**
Service	**0.18**	**0.44**

Since 2008-2009 forecast errors show major outliers and unusual discrepancies during financial recession, the period is not considered for the analysis of revision errors.

Table 3.5/B - Performance in real time: Sectoral Eurocoin

MACRO SECTORS	**Revision error**
SECTORS	**2003,1,2008,1**
Industry	0.05
Service	0.02

In terms of RMSFE the two aggregated macro indicators (concerning Industry and Services, respectively) give better results in terms of RMSFE than those concerning some specific sectors (Manufacturing, Energy, Construction, Energy, Financial). Perhaps, by aggregating data, volatility concerning bandpassed cycle decreases as much as real time performance improves in the sectoral model.

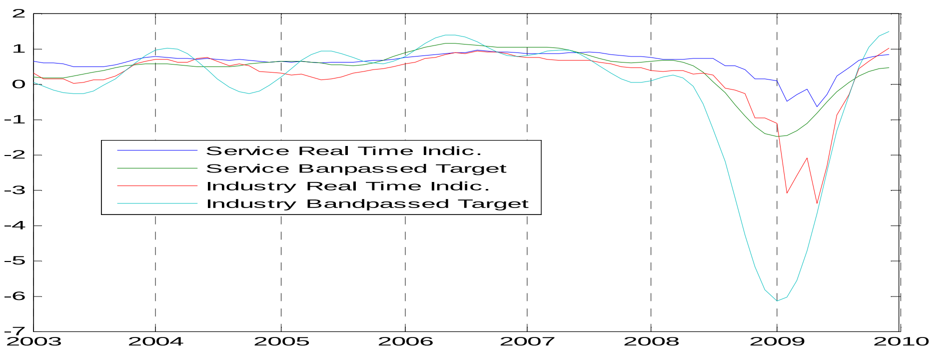
Concerning table 3.5/C below, we observe that PT two sided test is above 99% critical value (that is equal to 2.576) for Industry and for the Service aggregate. We can't accept the hypothesis that there is not relationship between the direction of change predicted by aggregate models and the observed change: the results are particularly efficient for the Service aggregate.

Table 3.5/C - Non-parametric Statistic of Pesaran-Timmermann (PT)

MACRO SECTORS	PT	p-value of the PT test statistic	% Correct prediction of sign of bandpassed target
Service	3.23	0.0012	0.67
Industry	2.76	0.0058	0.65

Also by a graphic analysis, see figure 3.8 below, we find that the performance in real time of the Service indicator versus its bandpassed target is better than that regarding "Industry "aggregate.

Figure 3.8 - Empirical Target and Real Time Indicators: Service versus Industry case



Concerning the behaviour of the aggregated indicators for Service and Industry, we observe the following results in terms of turning points (table 3.5/D-E and figures 3.9 below):

Table 3.5/D - Number of Turning Points in the Bandpassed Target from 2003 to 2009

SECTORS	TOTAL TURNING POINTS	DOWNTURNS	UPTURNS
Industry	4	2	2
Service	2	1	1

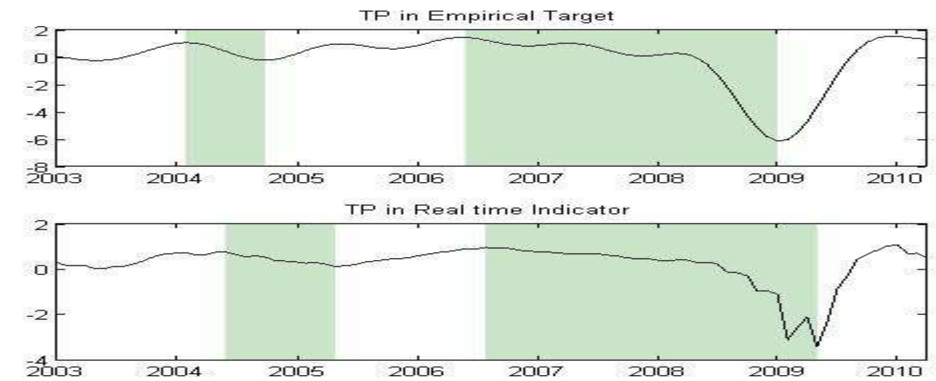
Table 3.5/E - Real time detection of turning points (TP)

AGGREGATED SECTORS	TP Signals	Correct TP	Correct over signalled TP	Missed over all TP
Industry	4	1	1/4	3/4
Service	4	1	1/4	1/2

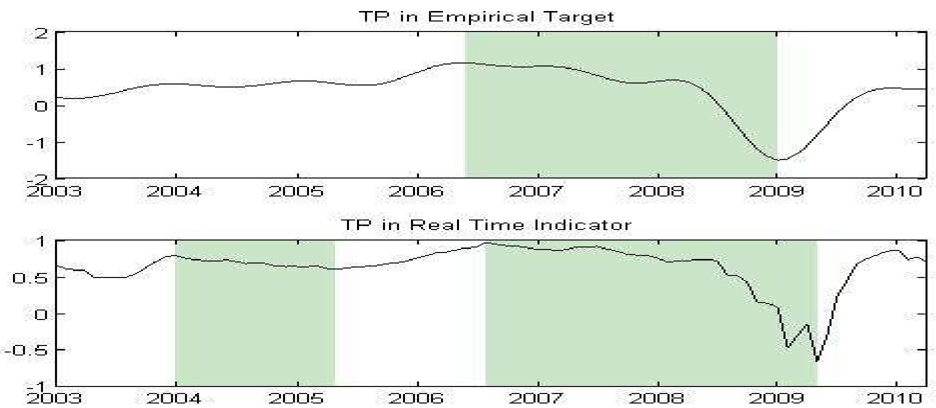
***Figures 3.9**: Real Time detection of Turning points in Sectoral Aggregates*

Shade Recession/ Contraction Area

A) Industry



B) Service

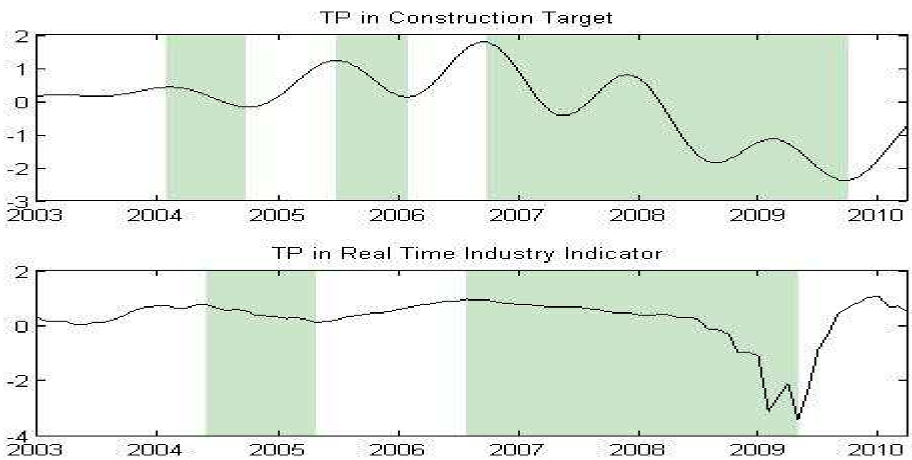


With regard to the cycle phases of each aggregated sectoral growth rate outlined above, we observe, particularly, that Service Real Time indicator is not able to signal the recession phase in 2004 (following Bry and Boschan procedure). In figure 3.11 below we outline ex post sectoral MLRG within the 1999-2011 period for the Service and Industry aggregate by our generalized dynamic factor models.

Finally, we can outline that if a simple index concerning Trade-Transport-Communication appears particularly efficient to assess the underlying direction in which this sectoral economy moves, it could be useful to assess the remaining sectoral growth rates (Manufacturing, Energy, Construction, and Financial) by using some information concerning macro area indicators[34]**.**
In fact, in figure 3.10 and 3.11 below, we observe that Real Time "Industry" Indicator is able to predict recession in "Construction" target business cycle in 2006, while Real Time "Service" index detects 2/4 of TP in "Financial" empirical target TP.

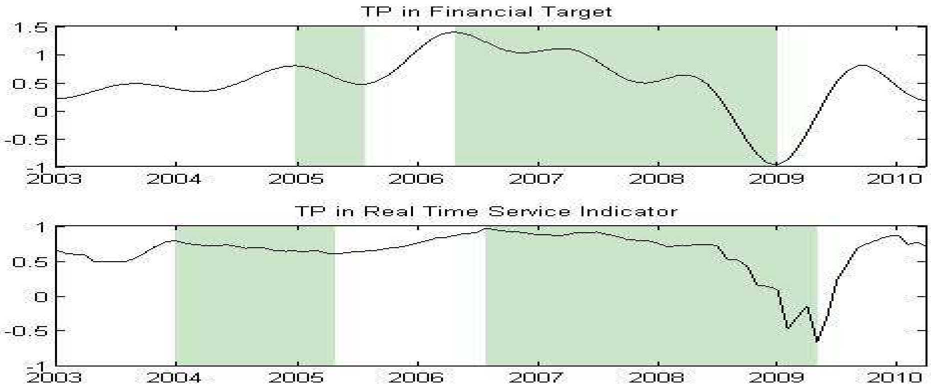
Shade Recession/Contraction Area

Figures 3.10: Construction Target versus Industry in Real Time



[34] By the development of generalized dynamic factor model, we can use "Industry Index" to forecast Manufacturing, Energy and Construction growth.

Figures 3.11: Financial Target versus Service in Real Time



While in this section we have tested our indicators for the main macro sectors with regard to Euro Area, in the next we will analyze performance in real time by a recursive window, with reference to European countries.

3.2.5 A Comparison with Bandpass Filters

In this sub-section we focus our analysis on Trade-Transport-Communication (TTC) indicator in real time, since it scores remarkably better than other sectoral indicators compared to the respective approximate ideal bandpass filter. We shortly show a systematic comparison of TTC, based on Eurocoin approach, with the Baxter-King (BK) and Christiano-Fitzgerald (CF) version of the band-pass filter. It seems useful to remember that the value of the bandpassed (filtered) data is available with good accuracy only at time T +h, for a suitable h. Therefore, our indicators produced at time T can be compared with the target at T produced at time T + h. We can therefore consider the Eurocoin approach as a method to forecast smoothed GDP, since we produce at time T sectoral preliminary estimates of the medium to long run component of the growth.

Table 3.6/A - Performance in terms of RMSE

Indicator	**RMSE with respect to c***	**RMSE with respect to c***
	2003,1,2008,1	**2003,1,2009,3**
TTC in Real Time	0.18	0.52
BK	0.62	0.63
CF	0.46	0.43

Table 3.6/B - Non-parametric Statistic of Pesaran - Timmermann (PT)

Indicator	PT	p-value of the PT test statistic	% Correct prediction of sign of bandpassed Δc^* 2003,1,2009,3
TTC in Real Time	2.16	0.0309	0.62
BK	5.15	0.0000	0.80
CF	1.62	0.1047	0.59

Table 3.6/C - Detection of turning points (TP)

Indicator	TP Signals	Correct TP	Correct over signalled TP	Missed over all TP
TTC in Real Time	**4**	**3**	**3/4**	**1/4**
BK	**3**	**2**	**2/3**	**2/4**
CF	**4**	**2**	**2/4**	**2/4**

We observe particularly that:

- regarding the nowcast of the sectoral target c_t^* and the ability to detect Turning Points, our Trade-Transport-Communication indicator (implementing Eurocoin methodology) scores remarkably better than CF and BK (particularly in 2003-2008). About the capacity to track the direction of the target, BK filter works better than the other indicators used.
- Furthermore, the Eurocoin approach gives a monthly frequency and timely estimate at the beginning of each month, differently from bandpass filters, released several months ahead of the other bandpassed estimates.

3.3 National Indicators: Real Time Experiment

In this section we will outline some national indicators explaining the medium to long run component (MLRG) of GDP by pseudo real time simulations.

Our goal is achieved by the generalized dynamic factor models for the following countries: Belgium, Italy, France, Germany, Spain, Netherlands (in chapter 2 we have already built these indicators by ex-post estimation).

Three different types of national indicators can be obtained by the strategies (methods) below indicated:

1) Projection of national GDP on European factors (so that we can use all the 157 national and European variables used by the Bank of Italy to build Eurocoin indicator);

2) projection of Euro Area GDP on factors built by national variables (e.g. projection of European GDP on the common factors that are linear combinations of the 23 French variables included in our Dataset);

3) projection of national GDP on common national factors (e.g. projection of French GDP on French common factors)".

Our elaborations are based in this research on the GDP by country expressed in constant prices. Data concerning national quarterly gross value added are available from Eurostat Statistics Database by themes.

Within the 2002-2010 finite sample we consider the following approximation of the target that can be obtained by augmenting the national growth rate y_t^{*i} for each country i, with its sample mean $\hat{\mu}$ in both infinite directions:

$$c_t^{*i} = \beta(L) y_t^{*i}, \text{ where } y_t^{*i} = \begin{cases} y_t & if \quad 1 \le t \le T \\ \hat{\mu} & if \quad t < 1 \;\; or\; t > T \end{cases} \tag{3.3}$$

Problem arises due to the fact that the national growth rate is observed only quarterly, while we are interested in a monthly indicator of economic activity, so that the interpolation is needed.

Our results could be influenced by the interrelation among European and National bandpassed growths (see table 3.7/A) and by the volatility in National bandpassed GDP (table 3.7/B), as we will be using both national and European variables to calculate common factors.

In the table below we show RMSE and correlation obtained by comparing National growth rate with Euro Area growth rate, smoothed by using a bandpass filter, see equation (3.3) above.

Table 3.7/A - **European Bandpassed GDP versus National Bandpassed GDP in 1995-2008**

	RMSE	**Correlation**
Italy	0.31	0.77
Belgium	0.27	0.70
France	0.16	0.84
Germany	0.26	0.87
Netherlands	0.31	0.75
Spain	0.39	0.77

Table 3.7/B – Volatility in National bandpassed GDP

COUNTRY	**BANDPASSED GDP 1995,6-2008,1**
Spain	**0.05**
France	**0.08**
Italy	**0.14**
Germany	**0.16**
Netherlands	**0.16**
Belgium	**0.15**
Euro Area	**0.08**

Among the countries considered in this exercise, we observe the highest national volatility in bandpassed growth for Belgium, Germany, Italy, Netherlands smoothed growth rate; the lowest for Spain and France (this is similar to the one concerning Euro Area). Germany MLRG is, from tables 3.7 A-B), the more similar to Euro Area MLRG.

In the following, we show for the considered countries results in a real time exercise concerning the three different approaches used to outline national indicators (we obtain three different indexes for each country).

Real time estimates are built both from 2003 to 2008 and from 2003 to 2010, separately, because in the 2008-2010 period, we observe a strong recession and an high variation in volatility concerning GDP. We establish the criteria to apply in pseudo-real time evaluation to analyze the ability of $\hat{c}_t^{\,i}(t)$ for each country i considered, to estimate (approximate) the national truncated band-pass filter $c_t^{*i}(T)$. Therefore, we analyze performance of the "National Eurocoin" at time t, with $t \leq T-12$, by the difference between our indicator at time t and the approximate target at t that is obtained using data up to T.

Monthly length of the sample 2002-2010 is equal to T=108 for the period T-96 ≤ t ≤ T-13. We are interested in:

a) the RMSE of nowcast errors concerning the 2003,1,2009,12, period is outlined by the root

of the ratio $\dfrac{\sum_{t=T-96}^{T-13}\left[\hat{c}_t^{\,i}(t)-c_t^{*i}(T)\right]^2}{84}$;

b) the size of the revision errors after one month.

c) the ability of $\hat{c}_t^{\,i}(t)$ - $\hat{c}_{t-1}^{\,i}(t)$ = $\Delta\hat{c}_t^{\,i}(t)$ to signal the correct change of the bandpassed variation $\Delta c_t^{*i}(T)$.

For every statistics outlined (points (a),(b),(c) above) and for each country analyzed, we show the following results.

Table 3.8 – RMSE among Real Time Performance and National Bandpassed GDP

	2003,1,2008,1			2003,1,2009,12		
COUNTRY	**Euro GDP on National Var.**	**National GDP on National Var.**	**National GDP on Euro Var**	**Euro GDP on National Var.**	**National GDP on National Var.**	**National GDP on Euro Var**
Italy	0.28	0.27	0.27	0.67	0.56	0.57
Belgium	0.29	0.31	0.30	0.59	0.68	0.51
France	0.19	0.17	0.12	0.44	0.37	0.32
Germany	0.37	0.38	0.39	0.69	0.63	0.62
Netherlands	------	-----	0.29	----	----	0.59
Spain	0.44	0.22	0.10	0.47	0.43	0.44

In table 3.8 above we consider the capacity of each national indicator in real time on nowcast of its bandpassed target. We observe that the projection of National GDP on European variables is generally the best strategy to adopt, followed by that of the projection of National GDP on National variables. These results tend to confirm the ones developed in section 2.3 for ex-post estimates concerning the 1995-2002 period.

In this section, it seems more useful and prudent to analyze performance in terms of RMSE from 2003 to 2008, before the unprecedented recession in 2008-2009 that strongly increases RMSE values.

Since for Netherlands the number of variables in the Dataset useful for building common factors is very low (only 5 variables), it is impossible to build some “Netherlands common factors”. Therefore the only strategy concerning the projection of National GDP on European variables is developed.

Table 3.9 below shows that:

- Italy and France Indicator performance is particularly good in terms of revision error by the three different projection strategies;

- For Germany, Netherlands and Spain it is useful to project National GDP on Euro variables to minimize the revision error;
- Belgium shows a bad performance concerning the size of the revision errors.

Table 3.9* – *Size of the revision errors after one month

	2003,1,2008,1		
	Euro GDP on National Var.	**National GDP on National Var.**	**National GDP on Euro Var**
Italy	0.030	0.033	0.026
Belgium	0.079	0.144	0.064
France	0.036	0.039	0.026
Germany	0.055	0.058	0.036
Netherlands	---	---	0.034
Spain	0.095	0.088	0.026

The null hypothesis of PT test ((see equation 1.37 and 1.39 in the annex 2) that is tested in tables 3.10 is that the real time indicator (the predictor) and the (predicted) bandpassed target are independently distributed, i.e. the national Eurocoin, for each country, has no power in nowcasting the target: the 95% and 99% critical values are, respectively, 1.96 and 2.576. PT test is developed for the January 2003 – December 2009 period.

Table 3.10/A

Non-parametric Statistic of PT: Strategy of the projection of National GDP on European variables

	PT	**p-value of the PT test statistic**	***% Correct prediction of sign of bandpassed*** Δ
Italy	3.47	0.0000	0.69
Belgium	0.57	0.5704	0.53
France	2.57	0.0103	0.64
Germany	3.16	0.0016	0.67
Netherlands	3.20	0.0014	0.67
Spain	1.86	0.0628	0.59

Table 3.10/B

Non-parametric Statistic of PT: Strategy of the projection of National GDP on National variables

	PT	**p-value of the PT test statistic**	***% Correct prediction of sign of bandpassed*** Δ
Italy	1.70	0.0895	0.58
Belgium	- 0.42	0.6752	0.48
France	0.98	0.3289	0.55
Germany	1.43	0.1517	0.58
Netherlands	---	--	---
Spain	0.31	0.7545	0.53

Table 3.10/C

Non-parametric Statistic of PT: Strategy of the projection of EURO GDP on National variables

	PT	**p-value of the PT test statistic**	***% Correct prediction of sign of bandpassed*** Δ
Italy	1.94	0.0524	0.60
Belgium	1.08	0.2807	0.57
France	1.87*	0.0608	0.60
Germany	1.38	0.1685	0.58
Netherlands	---	---	---
Spain	1.22	0.2212	0.58

* We obtain a PT equal to 2.15 for the period 2003,1-2010,12

By the projection of National GDP on Euro variables (table 3.10/A), we observe that PT two sided test is:

- above 99% critical value for Italy (particularly), Germany and Netherlands.
- above 95% critical value for France.

For Spain and Belgium we observe a lower performance in terms of correct prediction of sign; particularly for Belgium we observe that the significant level of the test is very low (very high p-value).

The projection of National GDP on National variables (table 3.10/B) shows a better performance for Italy and for Germany in terms of prediction of sign (0.58%).

By the projection of the European GDP on National variables (table 3.10/C), we observe that PT test is above 90% and near 95% critical value for Italy and France. For these countries we observe a sufficient prediction of signs in the bandpassed target.

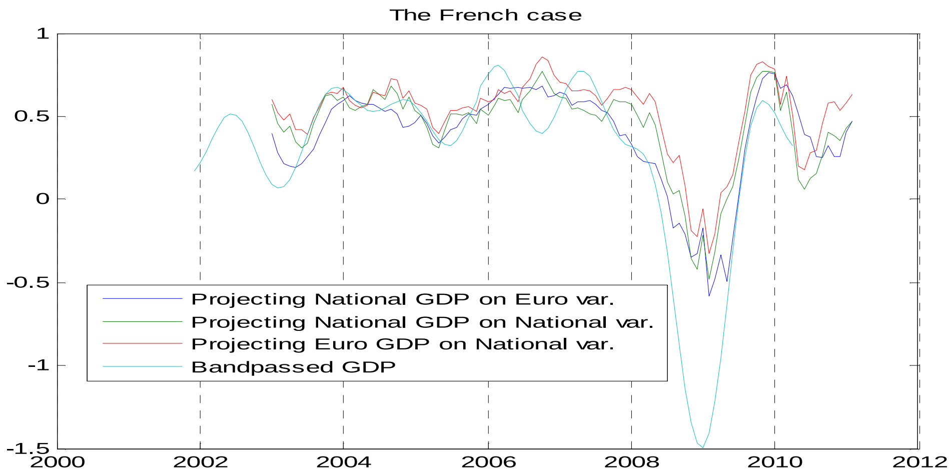
Some conclusions:

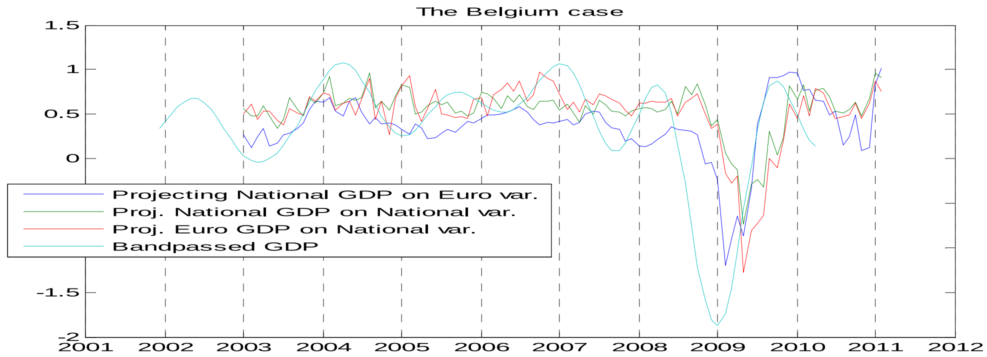
Analyzing real time performance in 2003-2008, we observe that:

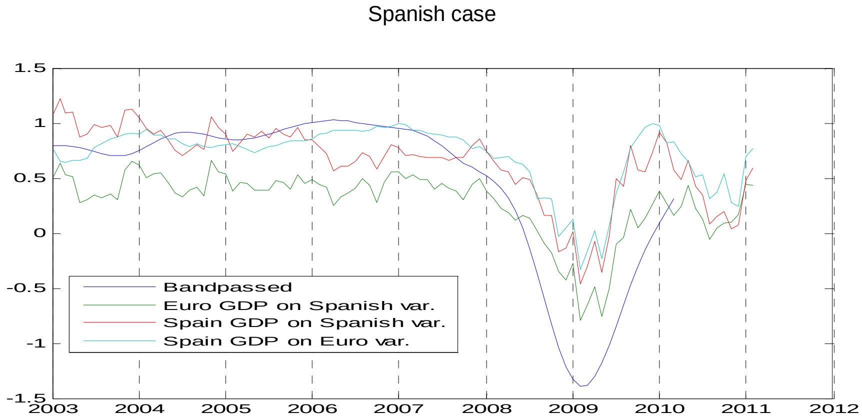
- in terms of RMSFE national indicators concerning Italy, France and Spain provide a fairly good approximation of band-pass GDP in real time, particularly by the projection of National GDP on the 157 European variables included in the dataset.
- During the 2008-2010 period, we observe a very high volatility that doesn't allow the obtaining of a clear comparison of our indicators. For France, RMSFE is particularly good, considering business cycle volatility (in this phase). The interrelations among a large dataset of European variables and National gross domestic product seem to better our estimates.
- Also, in terms of size of revision, Italy and France perform better than other sectors;
- concerning the ability of the real time indicator to track target direction, we show that Italy, France, Germany, Netherlands perform fairly well by the projection of National GDP on Euro variables.

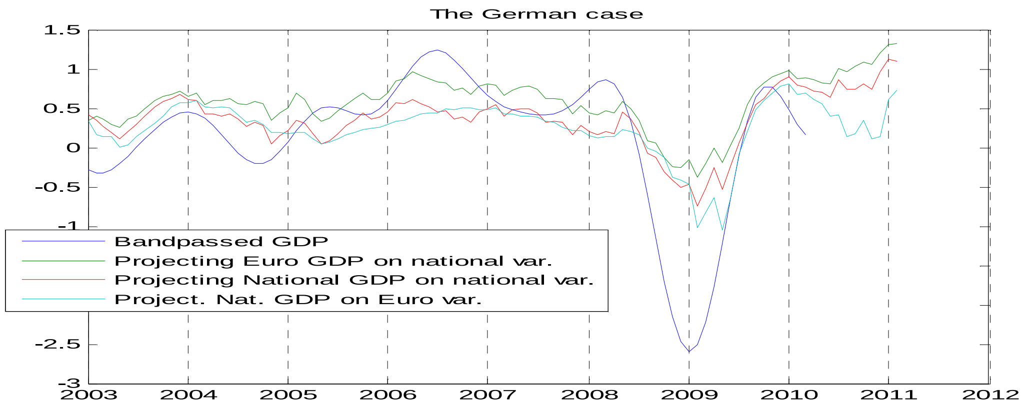
The figures below show, for each country, real time performance, obtained by using Eurocoin methodology and the three strategies outlined above.

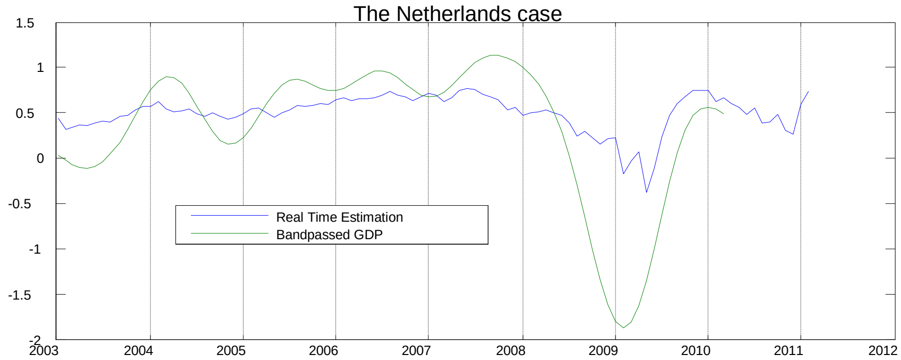
Figures 3.12 Comparing National Real Time Indicators with Bandpassed GDP



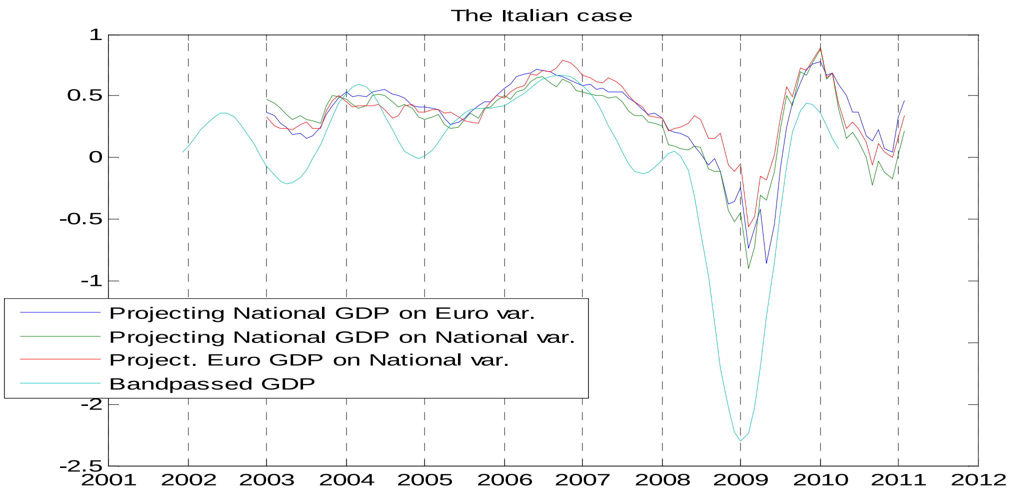

The Belgium case
Projecting National GDP on Euro var.
Proj. National GDP on National var.
Proj. Euro GDP on National var.
Bandpassed GDP


Spanish case
Bandpassed
Euro GDP on Spanish var.
Spain GDP on Spanish var.
Spain GDP on Euro var.


The German case
Bandpassed GDP
Projecting Euro GDP on national var.
Projecting National GDP on national var.
Project. Nat. GDP on Euro var.



It seems interesting to observe that real time estimation is generally not able to detect the depth of the recession in 2008-2010: e.g. for Italy (in the figure below) the minimum of the monthly quarter on quarter GDP growth rate in real time is equal to -0.9, while for the bandpassed target we have a minimum value of - 2.4.



3.3.1 Behaviour around Turning Points

A characteristic of these national indicators is the ability to give a correct signal of MLRG turning points (TP) in real time. We may say that an upturn (downturn) signal in $\hat{c}_t^s$ can be predicting or lagging true TP (i.e. a four-month error is tolerated). We outline the TP for each strategies considered in the national experiment by the Bry-Boschan procedure (see annex 1):

Case 1) Projection of National GDP on European Variables

According to this definition involved in the pseudo real-time exercise, the target exhibits the following turning points for each country in the 2003-2009 subsample:

Table 3.11/A - Number of Turning Points in the Bandpassed Target

COUNTRY	TOTAL TURNING POINTS	DOWNTURNS	UPTURNS
Italy	4	2	2
Belgium	4	2	2
France	4	2	2
Germany	4	2	2
Netherlands	4	2	2
Spain	5	3	2

Table 3.11/B - Real time detection of turning points (TP)

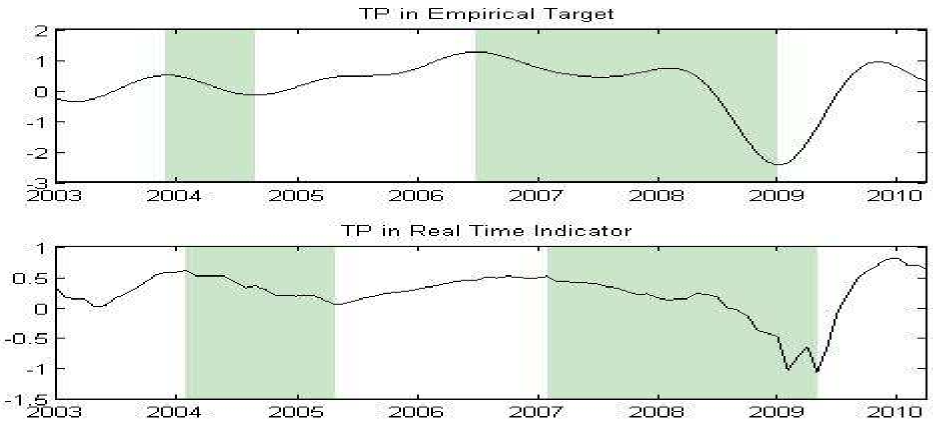
COUNTRY	TP Signals	Correct TP	Correct over signalled TP	Missed over all TP
Italy	4	3	3/4	1/4
Belgium	4	2	2/4	2/4
France	4	3	3/4	1/4
Germany	4	2	2/4	2/4
Netherlands	4	4	4/4	---
Spain	4	2	2/4	3/5

In the graphs below, the areas shaded green represent national recession/contraction, while the points at the edges of the areas shaded green represent the turning points of the cycle.

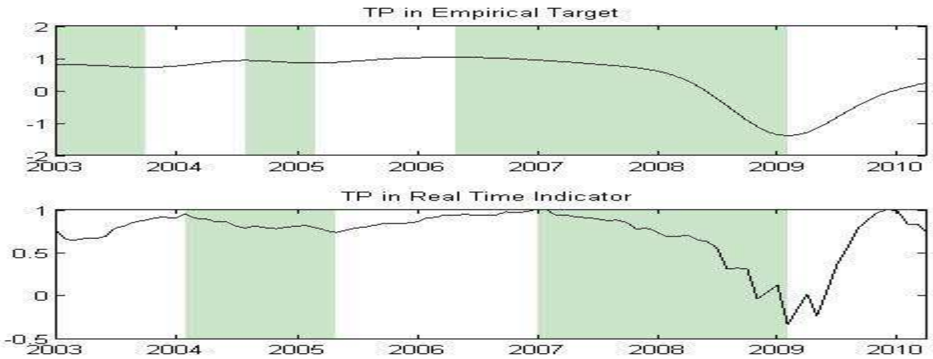
Figures 3.13: Real Time detection of National Turning points

Shade Recession Area

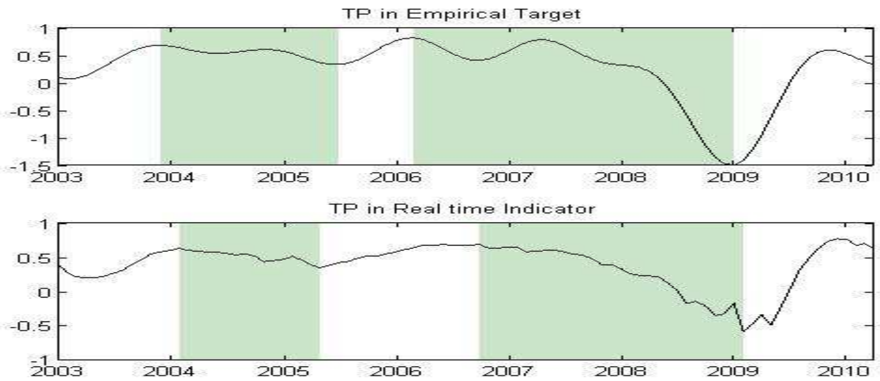
A) GERMANY



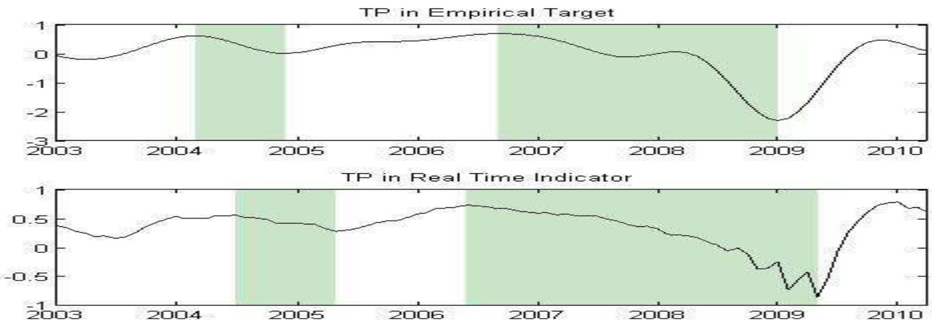
B) SPAIN



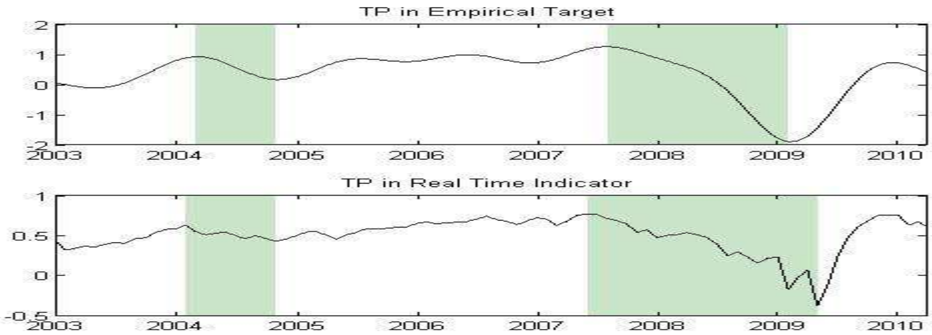
C) FRANCE



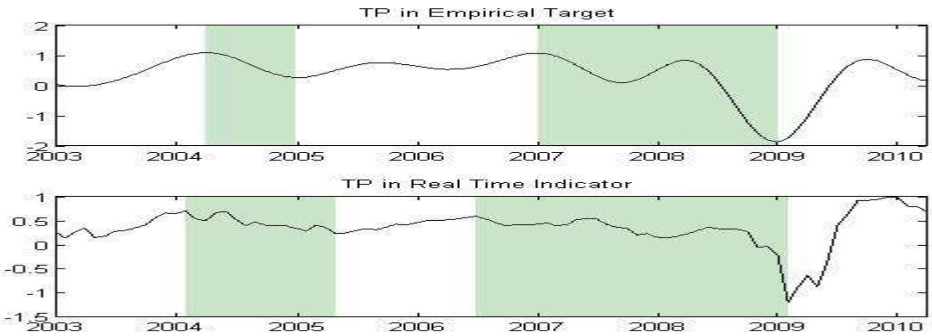
D) ITALY



E) NETHERLANDS



F) BELGIUM



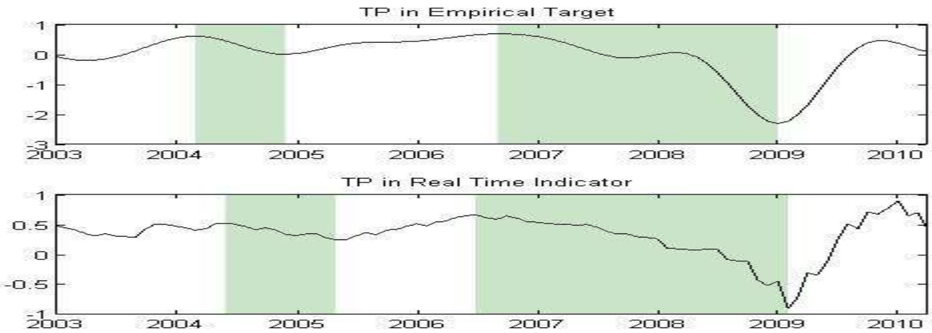
Concerning the cycle phases outlined above, we observe that:

- the Belgium and Netherland indicators are always able, in the considered sample, to predict recession in the target;
- the capacity of TP in Italian Real Time to predict the recession phase in 2006 in relation to Empirical Target, while Spanish indicator perform poorly in signalling cycle phase.

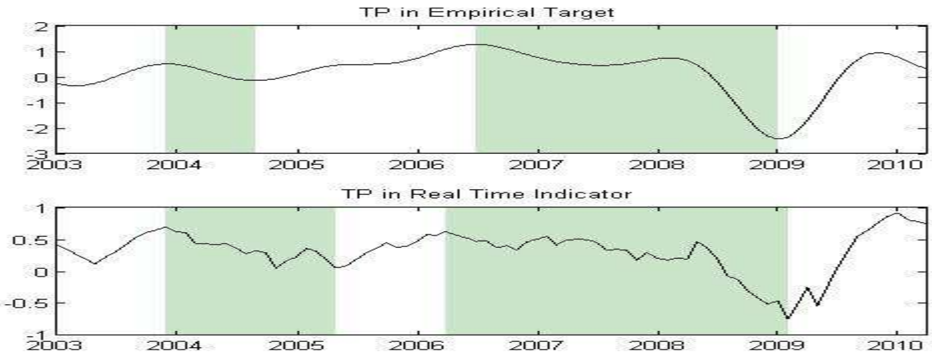
We have developed above in ***Case 1)*** TP for the most performing national strategy (that concerning the projection of National GDP on a large dataset of European variables), among the three considered. Analyzing results of test and statistics developed above, we decide to build Turning Points and Cycle phases by the remaining strategies only for the following countries (the main criteria that we have used is to include country-indicator if the % of correct prediction of signs with regard to the bandpassed target is at least equal to 0.58%.):

Case 2) Strategy of the projection of National GDP on National variables:

a) ITALY



b) GERMANY

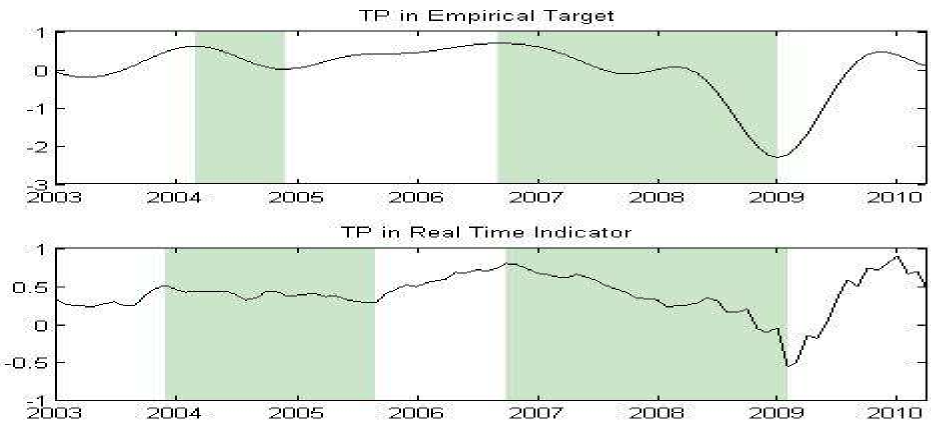


Concerning the cycle phases outlined above in **Case 2)** we observe the capacity of TP in Italian and German real time performance to predict the recession phase in 2006 in relation to Empirical Target. Both for Germany and for Italy, we observe the capacity to signal, by this strategy, 3/4 of the turning points in the empirical target.

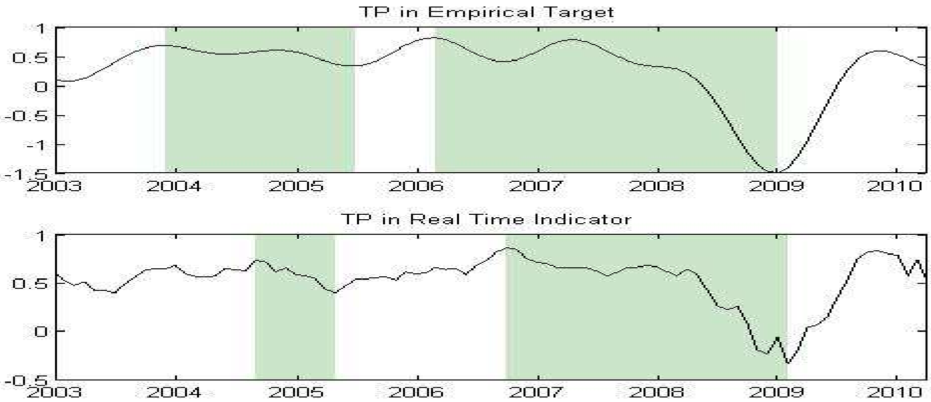
Case 3): Strategy of the projection of Euro GDP on National variables:

We analyze this strategy only for correct country-indicators predicting at least 60% of the signs in bandpassed empirical target.

a) ITALY



b) FRANCE



Concerning the cycle phases outlined above in **Case 3)** we observe TP capacity in Italian Real Time Indicator, obtained by projecting Euro GDP on the 22 Italian variables in Thomson Financial Datastream used in this research to predict the 2004 Italian recession phase and to signal 3/4 of the turning points in the empirical target. Perhaps, relevant interrelations between the Italian and European economies can produce these results. On the contrary, the French indicator performs very poorly in signalling cycle phase

3.3.2 A Comparison with Bandpass Filters

In this sub-section we focus on the analysis of Italian indicator in real time (obtained by the projection of the Italian GDP on the national common factors) based on generalized dynamic factor model (GDFM), since it scores remarkably better than the other national indicators, compared to the respective approximate ideal bandpass filter. We shortly show a systematic comparison of TTC, based on Eurocoin approach, with the Baxter-King (BK) and Christiano-Fitzgerald (CF) version of the band-pass filter.

Table 3.11/C - Performance in terms of RMSE

Indicator	**RMSE with respect to c***	**RMSE with respect to c***
	2003,1,2008,1	**2003,1,2009,3**
GDFM	0.27	0.58
BK	0.31	0.46
CF	0.26	0.29

Table 3.11/D - Non-parametric Statistic of Pesaran - Timmermann (PT)

Indicator	PT	p-value of the PT test statistic	% Correct prediction of sign of bandpassed Δc^* 2003,1,2009,3
GDFM	3.12	0.0018	0.68
BK	6.33	0.0000	0.86
CF	5.17	0.0000	0.79

Table 3.11/E - Detection of turning points (TP)

Indicator	TP Signals	Correct TP	Correct over signalled TP	Missed over all TP
GDFM	**4**	**3**	**3/4**	**1/4**
BK	**5**	**3**	**3/5**	**1/4**
CF	**4**	**4**	**4/4**	**-**

We observe particularly that:

- regarding the nowcast of the sectoral target c_t^*, our Italian indicator based on GDFM (implementing Eurocoin methodology) scores remarkably better than BK. About the capacity to track the direction of the target, BK filter works better than the other indicators used. With regard to the ability to detect Turning Points, CF filter gives a better performance.
- However, the Eurocoin approach gives for Italy a monthly frequency and timely estimate at the beginning of each month, differently from bandpass filters, released several months ahead of the other bandpassed estimates.

3.3.3 Some Final Remarks

By the statistics developed in this section, it is clear that by applying Eurocoin methodology based on dynamic factor model and by using European and National variables contained in Thomson Financial Datastream (TFD), for Italy, France and Netherlands we can obtain some national indicators with a good performance that can compete with the truncated band-pass filters. In addition, they do not deteriorate at the end of the sample (see figures 3.12 above). The strategy used in obtaining these national performances is projecting National GDP on the European common factors that are a linear combination of the 157 European variables included in the dataset.

However, we obtain interesting results for Italy and Germany by projecting National GDP on National common factors (e.g. by the projection of German GDP on a linear combination of the *observable* German monthly data in the TFD).

In synthesis, it may be possible to estimate the month-on-month changes of the smoothed GDP (*unobservable*) both by using a large dataset with 157 variables in the case of the TFD and also by a more parsimonious number of variables.

3.4. Combining Real and Financial Variables in Real Time

In this section we briefly show the performance in pseudo real time by trying to improve Eurocoin methodology (see par.2.4, chapter 2) dividing the 157 variables contained in Thomson Financial Datastream in real and financial variables (econometric techniques to combine the forecasts are presented in chapter 1). A subdivision among real and financial economies is substantially confirmed by Forni et al. (2003), where they show that these financial variables do help in forecasting inflation, but do not in forecasting industrial production.

Our experiment is useful in the analysis of the impact of real data on estimate smoothed GDP in the different business cycle phases. So, we have outlined two data groups: a first containing "*real economic activity variables*" and a second with "*financial variables*".

In par. 2.4 the combination of real and financial MLRG (by using the regression method to determine the relative weights) is called "*Combined Eurocoin*". A synthesis of different approaches to combine more forecasts is described in chapter 1, par.1.6.

Two groups of common factors are used to project GDP: F_i and R_i ($i = 1,...m$) will be the common factors relevant to the prediction of "real MLRG" and "financial MLRG" , obtained by projecting Euro Area GDP respectively on real and financial variables. The weights used to combine the two forecasts will be obtained by the regression method.

In the tables below, results of an exercise are shown, concerning correlations (and Rmse) of the following indicators versus European bandpassed GDP (the target):

- Real Eurocoin (named "Real variables" in figure 3.14)
- Financial Eurocoin (named "Financial variables" in figure 3.14)
- Combined Indicator
- Eurocoin.

The 1995-2010 period is investigated in this thesis to analyze business cycles. Real time estimations are built from 2003 to 2008 and from 2003 to 2010 separately, as in 2008-2010 we observe a strong recession and an high variation in GDP volatility.

In this exercise our approximate target is the bandpassed Euro Area growth rate, the same that is generally used to test the performance of the Eurocoin indicator. Our real time indicator will be a combination of real and financial variables by using the weights produced by a simple regression model, as we will explain below.

In table 3.12/A we test the capacity of the "combined estimation" inside the sample and in real time on nowcast the bandpassed target.

Table 3.12/A – RMSFE among Indicators and European bandpassed GDP

	1995,6,2002,12 (Rmse within the sample)	2003,1,2008,1 (Rmse in Real Time)	2003,1,2010,12 (Rmse in Real Time)
Real Eurocoin	**0.12**	**0.129**	**0.44**
Financial Eurocoin	**0.14**	**0.161**	**0.52**
Eurocoin	**0.12**	**0.125**	**0.43**
Combined Indicator	**0.12**	**0.107**	**0.39**

Source: Own elaboration using Thomson Financial Datastream

In table 3.12/B we analyze correlation between the "Combined Eurocoin" and the bandpassed target, both inside the sample and in real time. We observe that performance of Eurocoin Indicator is strongly similar to the one concerning the Real Indicator.

Table 3.12/B – Correlation among Indicators and European bandpassed GDP

	1995,6,2002,12 (Within the sample)	2003,1,2008,1 (In Real Time)	2003,1,2010,12 (In Real Time)
Real Eurocoin	**0.91**	**0.87**	**0.88**
Financial Eurocoin	**0.89**	**0.77**	**0.77**
Eurocoin	**0.92**	**0.88**	**0.89**
Combined Indicator	**0.97**	**0.90**	**0.89**

Source: Own elaboration using Thomson Financial Datastream

Real time estimation shows that in 2003-2008 a combination between "real MLRG" and "financial MLRG" obtained projecting Euro Area GDP on real and financial variables respectively, can compete with the truncated band-pass filters within the sample and it can slightly outperform Eurocoin in terms of correlation (figure 3.14) and RMSE (figure 3.16), particularly in 2005-2007.

Figure 3.14: Pseudo Real Time Estimation

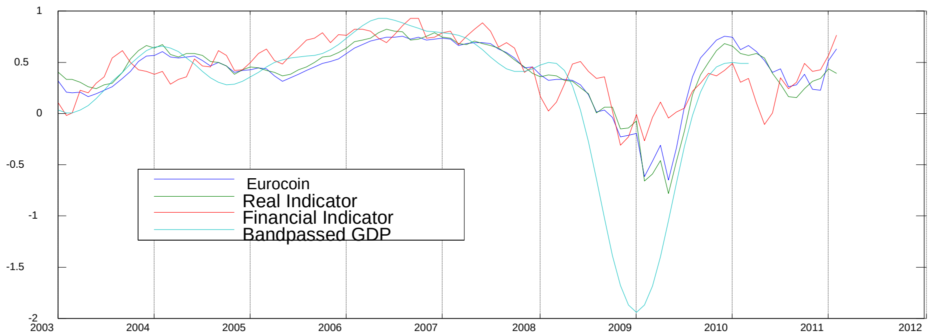
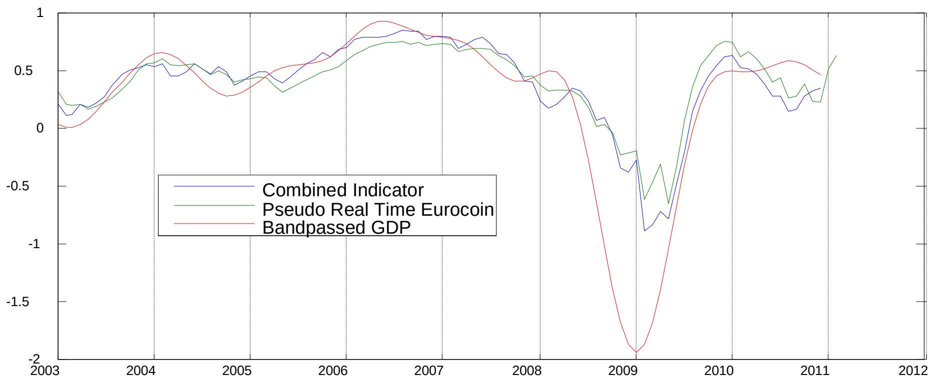
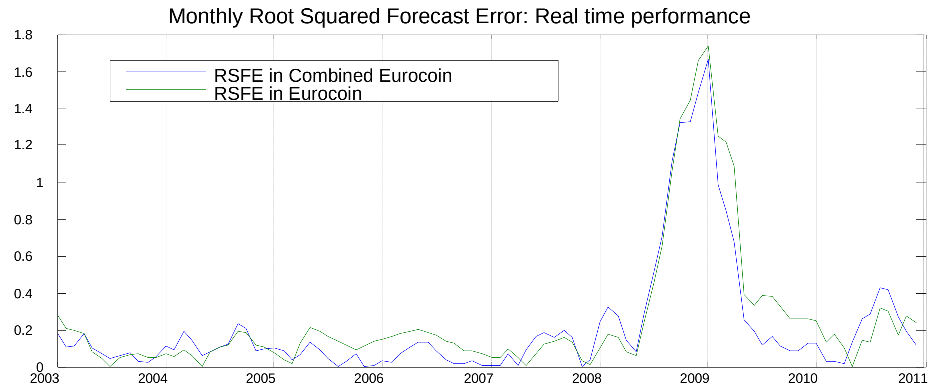


Figure 3.15 : Combined Eurocoin versus Eurocoin in real time



In figure 3.16 below RSFE is obtained by comparing Combined Eurocoin and Eurocoin in real time respectively to the bandpassed Euro Area GDP: in 2005-2007, regarding nowcast of bandpassed target, the "Combined Eurocoin" scores better than the classic Eurocoin.

Figure 3.16



In the equation (3.4) below the weights α and β are shown updated monthly to underline our regression in real time estimation period (2003-2010)

$$c_t = \alpha_t + \beta_{1t} c_t^R + \beta_{2t} c_t^F$$

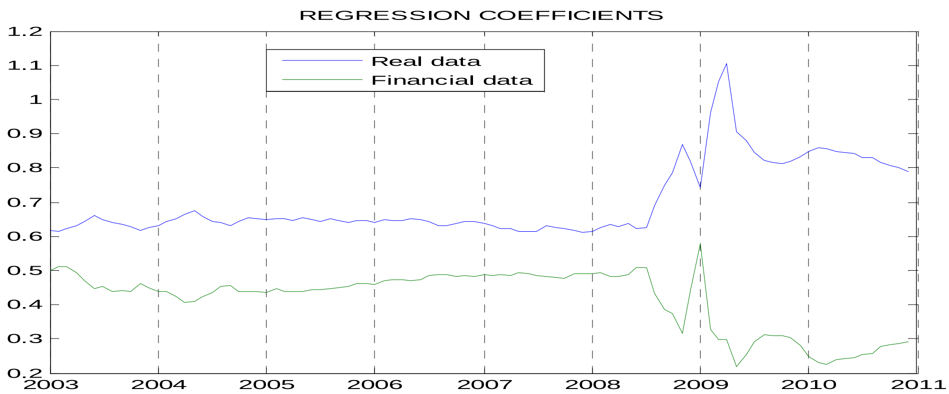
(3.4)

in which c_t indicates bandpassed GDP; c_t^R is the "Real Eurocoin" indicator that we calculate only using real variables; c_t^F is the "Financial Eurocoin" indicator that we calculate using financial variables only. The weights are updated every month on the basis of the newly available information.

This section focuses on the pseudo real time estimating performance of these disaggregated indicators. It is worthwhile to specify that the term *"pseudo"* indicates that the matrix of data we use does not consider the revision to the series. The same series are cropped over and again to take account of restriction to the set of information. Furthermore, we make the hypothesis, calculating weights in regression models (3.4), bandpassed GDP in real time are known. Our exercise is useful to asses the impact of real data to estimate MLRG for the different phases of the cycle. We use these weights to build our "Combined Eurocoin".

In figure 3.17 that follows we show the weights (regression coefficients) calculated in the combination of real and financial indicator. We observe that the relation between the two coefficients is quite stable till 2008; at the beginning of the last recession (during 2008), it changes the impact of real and financial data to estimate smoothed GDP, and in 2009 (at the trough) the distance becomes the minimum; during 2009, when recovery begins, we observe that the impact of real data to estimate GDP becomes more important than the one concerning financial data.

Figure 3.17



We can assess that the impact of real and financial variables in estimating smoothed GDP, during the structural break in 2008, shows that the role of real data as industrial production, demand indicators, foreign trade (Import, Export), Employment Indexes, becomes particularly relevant in relation to that concerning financial data as Exchange rates, Money Supply, Spreads. So, one possible explanation could be that interrelations among the recession phase and the variations in production, consumptions and unemployment are highly correlated.

Clearly, such conclusions are not general but restricted to the data and the models used in this exercise.

3.5 Smoothed GDP Components: Testing the Performance

In this paragraph we test a framework for smoothed GDP expenditure components (following the methodology built in chapter 2, par.2.5, by generalized dynamic factor model) by pseudo real time estimation. We will test the projection of Euro Area GDP components (Consumptions, Investment, Foreign Trade) on European factors.

The variables that we project are:

7) Final Consumption Expenditure;
8) Household Consumptions;
9) Gross Capital Formation;
10) Exports of goods and services;
11) Imports of goods and services.

We compare our estimates in *real time* with *in sample* performance (developed in chapter 2) and with *bandpassed data.* Data concerning quarterly gross value added by components are available from Eurostat Statistics Database by themes.

Within the 2002-2010 finite sample we consider the same approximation of the target outlined in equation (3.2), that can be obtained by augmenting the GDP expenditure components growth rate y_t^{*z} , for each component z, with its sample mean $\hat{\mu}$ in both infinite directions:

A problem arises due to the fact that the growth rate is observed only quarterly, while we are interested in a monthly indicator for each expenditure component, so that the interpolation is needed.

We analyze the ability of the real time indicator $\hat{c}_t^{z}$ for each expenditure component z , with $z = 1,..5,$ in estimating the approximate truncated band-pass filter $c_t^{*z}(T)$. Therefore, we analyze performance at time t, with $t \leq T - 12$, by the difference between our indicator at time t and the approximate target at t that is obtained using data up to T.

Monthly length of the 2002-2010 sample is equal to T=108. We are interested, for the period T-96 ≤ t ≤ T-13, in:

a) the RMSE of nowcast errors concerning the 2003-2009 period, is outlined by the root of the ratio $\dfrac{\sum_{t=T-96}^{T-13}\left[\hat{c}_t^{z}(t) - c_t^{*z}(T)\right]^2}{84}$;
b) the size of the revision errors after one month in terms of RMSE.

c) the ability of $\hat{c}_t^{\,z}(t) - \hat{c}_{t-1}^{\,z}(t) = \Delta\hat{c}_t^{\,z}(t)$ to signal the correct change of the bandpassed variation $\Delta c_t^{*z}(T)$.

In tables 3.13 we analyze the behaviour of real time estimates following the points (a),(b),(c) outlined above.

Table 3.13/A - Performance in real time for Expenditure Components of GDP

VARIABLES	2003,1,2008,1 Rmse with respect to c*	2003,1,2009,12 Rmse with respect to c*
Final Consumptions	0.10	0.20
Household Cons.	0.16	0.25
Investments	0.49	-----
Exports	0.57	1.49
Imports	0.49	1.21

Note: Sample January 2002-December 2010; the sub-sample in the real time exercise is 2003-2009

Table 3.13/B - Performance in real time: Revision errors after one month

VARIABLES	RMSE revision errors 2003,1,2008,1
Final Consumptions	0.014
Household Cons.	0.019
Investments	0.115
Exports	0.131
Imports	0.090

Table 3.13/C - Non-parametric Statistic of Pesaran – Timmermann (PT)

	p-value of the PT test statistic	PT	% Correct prediction of sign of bandpassed Δc* 2003,1,2009,12
Final Consumptions	0.0027	3.00	0.66
Household Cons.	0.0367	2.08	0.61
Investments	0.8575	0.18	0.51
Exports	0.0252	2.24	0.61
Imports	0.0012	3.24	0.67

By the statistics produced above we observe that:

1) regarding the capacity of each Euro Area GDP component on *nowcast of its bandpassed target*, the unprecedented recession of 2008-2009 increases strongly RMSE values. In terms of RMSE with respect to bandpassed data both Final Consumptions and Household Consumptions have been particularly good in estimating growth in the recent past. Furthermore, for Consumptions in sample performance developed in chapter 2, section 2.5, it is very similar to that concerning real time estimation. However, estimates concerning government expenditure are relatively bad in terms of correlation with bandpassed components. Regarding foreign trade (exports, imports) we observe the goodness of our pseudo estimates in terms of correlation, but RMSE appears particularly high (both for in sample estimates and out of sample);

- also for Investment, RMSE appears quite high.

2) in terms of *size of revision errors*, Final and Household Consumptions perform better than other components;

3) concerning the ability of real time indicators to *track the direction of the target,* we observe that Final Consumption and Import indicators perform well. PT test is above 2.33, 99% critical value for a two sided test, whereas for Household Consumptions and Export we have a good % of sign prediction but PT is only above 95% critical value. We show a bad performance for Investment variable.

Thus, the 157 European variables from Thomson Financial Datastream (TFD), that we use in building our common factors are particularly useful in building consumption indicators, but their whole contribution is less clear for foreign trade. The figures below show, for each GDP component that we develop through our models, pseudo real time estimates and their comparisons with the relative banpassed components.

Estimation in real time of smoothed Final Consumption growth (figure 3.18) include the following variables:

a) Household Consumptions;

b) Government Expenditure.

Figure 3.18 – Final Consumptions

Figure 3.19 – Household Consumptions

Estimation in real time of smoothed Investments growth (figure 3.20) include the following variables:

1) Constructions Sector growth;
2) Investments different from Constructions.

Figure 3.20 - Investments

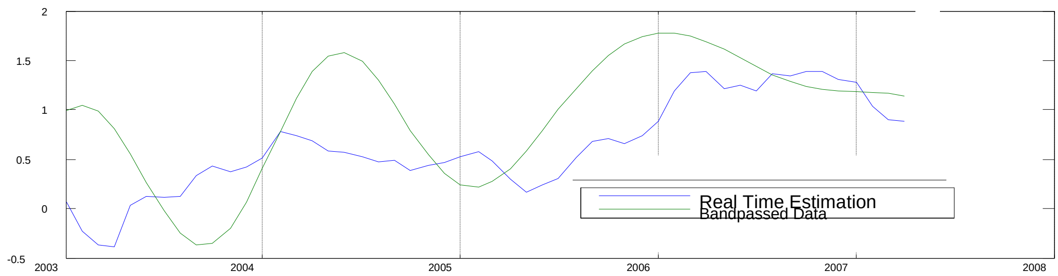


Figure 3.21- Exports

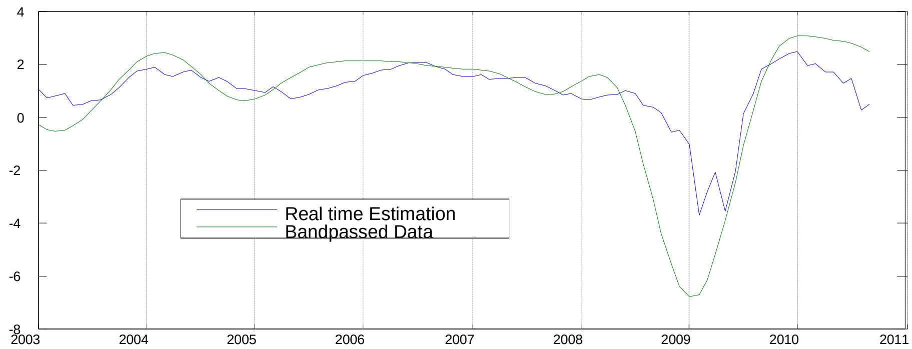
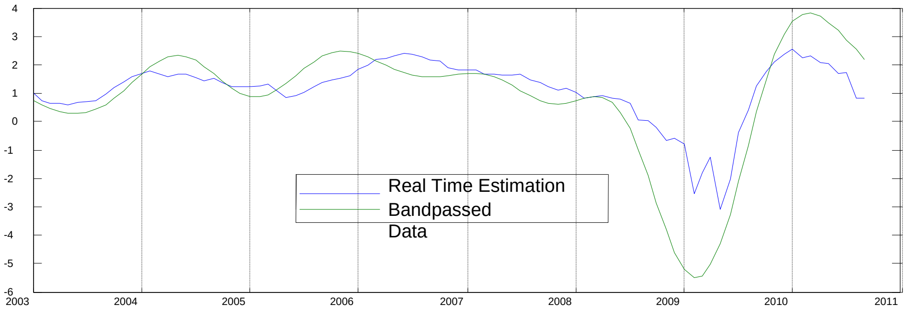


Figure 3.22 - Imports



In figure 3.21 and 3.22 above, we observe that it is impossible to intercept the depth of 2009's recession in real time. We also tested to project Export data only on the 11 variables that regard Foreign Trade and contained in TFD; no improvement in terms of RMSFE is obtained and correlation decreases at 0,46 during 2003-2008, against 0,84 by using all the 157 variables available in Datastream.

3.5.1 Behaviour around Turning Points

In paragraph 3.2.3 we have defined a *turning point* as a slope sign in the target, that is $c_t^{*z}(T)$ for each expenditure component z , assessing a downturn or an upturn. We may say that a real time upturn (downturn) signal in $\hat{c}_t^{s}$ can be predicting or lagging true upturn (therefore a four-month error is tolerated). We will outline turning points (TP) by the Bry-Boschan procedure.

According to this definition involved in the pseudo real-time exercise, in the 2003-2010 subsample, for each component of GDP, the target exhibits the following TP:

Table 3.14 – Number of Turning Points in the Bandpassed Target

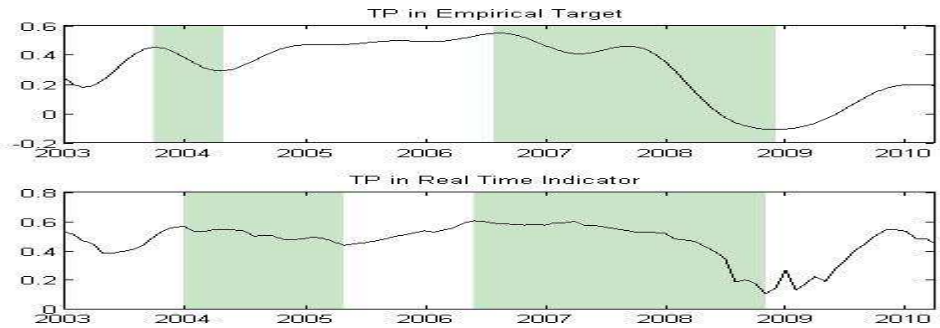
EXPENDITURE COMPONENTS	TOTAL TURNING POINTS	DOWNTURNS	UPTURNS
Final Consumptions	**4**	**2**	**2**
Household Cons.	**4**	**2**	**2**
Investments	**4**	**2**	**2**
Exports	**4**	**2**	**2**
Imports	**4**	**2**	**2**

Table 3.15/A: Real time detection of turning points (TP)

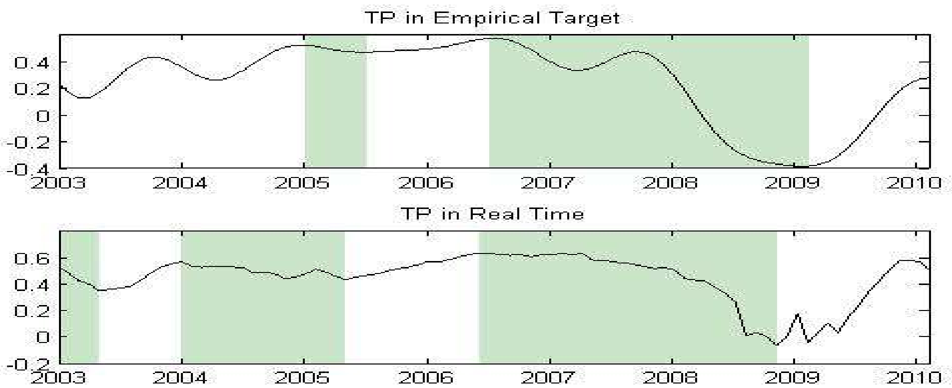
EXPENDITURE COMPONENTS	TP Signals	Correct TP	Correct over signalled TP	Missed over all TP
Final Consumptions	**4**	**3**	**3/4**	**1/4**
Household Cons.	**5**	**3**	**3/5**	**1/4**
Investments	**3**	**1**	**1/3**	**3/4**
Exports	**4**	**2**	**2/4**	**2/4**
Imports	**4**	**3**	**3/4**	**1/4**

Figures 3.23: Real Time detection of Turning points in Expenditure Components

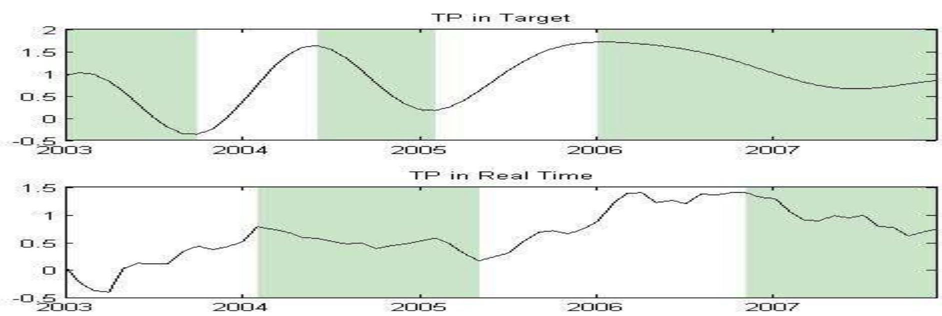
A) FINAL CONSUMPTION EXPENDITURE



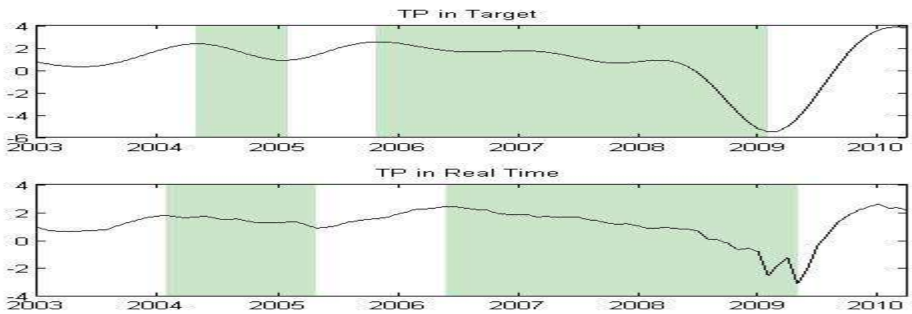
B) HOUSEHOLD CONSUMPTION EXPENDITURE



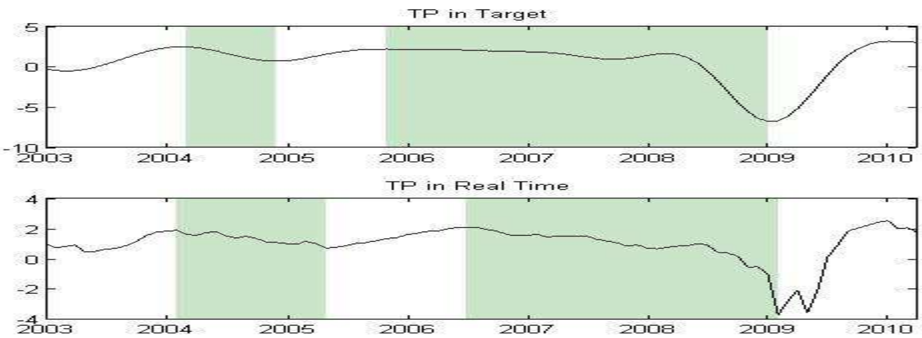
C) GROSS CAPITAL FORMATION (INVESTMENTS)



D) IMPORTS



E) EXPORTS



From the analysis of cycle phases we observe TP capacity in real time in Final and Household Consumptions to predict the 2006 recession phase, whereas Import and Export indicators are able to predict a contraction of the economy in 2004. Finally, we can assess both Final Consumption and Export indicators that we have outlined as efficient estimates of the respective medium to long-run components, performing equally well within and at the end of the sample.

3.5.1.1 A Comparison with Bandpass Filters

In this sub-section our analysis is focused on Final Consumption indicator in real time based on generalized dynamic factor model (GDFM), since it scores remarkably better than the other expenditure component indicators that we have analyzed, compared to the respective approximate ideal bandpass filter. We shortly show a systematic comparison of TTC, based on Eurocoin approach, with the Baxter-King (BK) and Christiano-Fitzgerald (CF) version of the band-pass filter.

Table 3.15/B - Performance in terms of RMSE

Indicator	RMSE with respect to c*	RMSE with respect to c*
	2003,1,2008,1	**2003,1,2009,3**
GDFM	0.10	0.15
BK	0.41	0.38
CF	0.35	0.33

Table 3.15/C - Non-parametric Statistic of Pesaran - Timmermann (PT)

Indicator	PT	p-value of the PT test statistic	% Correct prediction of sign of bandpassed Δc^* 2003,1,2009,3
GDFM	2.51	0.0120	0.65
BK	3.24	0.0012	0.67
CF	2.06	0.0391	0.62

Table 3.15/D - Detection of turning points (TP)

Indicator	TP Signals	Correct TP	Correct over signalled TP	Missed over all TP
GDFM	**4**	**3**	**3/4**	**1/4**
BK	**6**	**2**	**2/4**	**2/4**
CF	**3**	**1**	**1/3**	**3/4**

We observe particularly that:

- our Final Consumptions indicator (based on Eurocoin methodology) scores remarkably better than CF and BK, in terms of RMSE and with regard to the capacity to detect TP. About the capacity to track the direction of the target we observe a similar performance among the three indicators analyzed. Furthermore, the Eurocoin approach gives for Final

Consumptions a monthly frequency and timely estimate at the beginning of each month, differently from bandpass filters, released several months ahead of the other bandpassed estimates.

3.5.2 Estimating National Household Consumptions

In this section we outline national indicators explaining the medium to long run component of the growth (MLRG) for expenditure components of GDP. Variable projected is Household Consumptions. Both by the Projection of National GDP components on European factors and projecting National GDP components on National Factors, we estimate smoothed Consumptions. We compare the Italian case to the one concerning Germany. In tables 3.16 and 3.17, performance concerning these indicators is assessed in real time.

Table 3.16 - The Italian Case

Strategies	2003,1,2008,1 Rmse with respect to c* (Real Time Estimates)	RMSE revision errors 2003,1,2008,1	% Correct prediction of sign of bandpassed Δc* 2003,1,2009,12
Projection of National Household Consumptions on European Variables	0.23 (0.25)	0.044 (0.076)	0.52
Projection of National Household Consumptions on National Variables	0.16 (0.20)	0.039 (0.071)	0.59

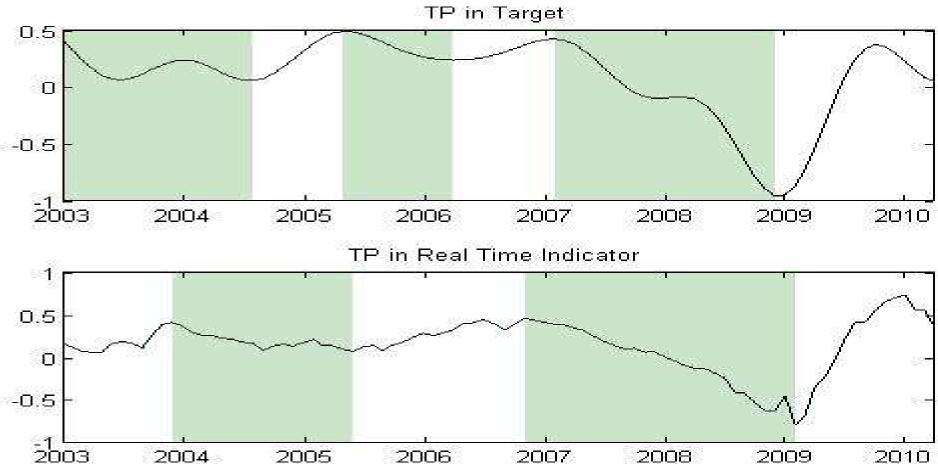
*** Statistics in parenthesis regard the 2003-2010 period.**

Table 3.17 The German Case

Strategies	2003,1,2008,1 Rmse with respect to c* (Real Time Estimates)	RMSE revision errors 2003,1,2008,1	% Correct prediction of sign of bandpassed Δc* 2003,1,2009,12
Projection of National Household Consumptions on European Variables	0.38 (0.51)	0.032 (0.052)	0.41
Projection of National Household Consumptions on National Variables	0.38 (0.55)	0.061 (0.076)	0.51

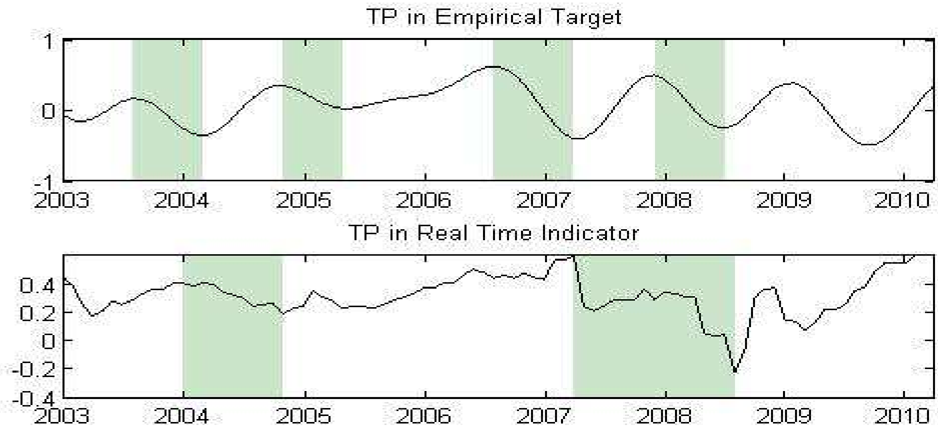
*** Statistics in parenthesis regard the 2003-2010 period.**

Figure 3.24 Real Time detection of Turning points in Italian Consumptions



From the results assessed in tables 3.16 – 3.17, the projection of National Consumptions on national variables (respectively for Germany and Italy) is the best strategy to be adopted.

Figure 3.25 Real Time detection of Turning points in German Consumptions



In this sub-section two different methods have been used to project bandpassed expenditure components of the growth through generalized dynamic factor models. From results obtained, our main findings can be summarized as follows:

a) in terms of RMSE and sign prediction, a national indicator in real time concerning Italy provides a good approximation of Italian band-pass Consumptions, particularly by the projection of National Consumptions on the 22 Italian variables included in the dataset. This strategy can maximize the capacity of our indicator to estimate Household Consumptions MLRG at the end of the sample;
b) about German performance, the choice regarding the best strategy to be adopted is less clear.

In synthesis it appears interesting to highlight that the smoothed estimates included in section 3.3 concerning National Eurocoin show how it is particularly useful to apply European variables to estimate national smoothed growth rate by dynamic factor model. Differently, to estimate National Household Consumptions (section 3.5.2), the impact of European (non national) variables seems less relevant.

Clearly, such conclusions are not general but restricted to the data and the models used in this exercise.

Chapter 4: Conclusions and Further Developments

Eurocoin indicator, published monthly by the Banca d'Italia and CEPR, provides a summary index of the medium to long-run component (MLRG) of GDP only for the whole Euro area aggregate. Therefore it is an aggregate indicator.

The innovation of this thesis are some disaggregated procedures based on Eurocoin approach, to give a monthly smoothed estimate of quarter-on-quarter growth rate with regard to: Euro Area sectoral data, some European countries, smoothed GDP components (consumption, investment, exports, imports) concerning Euro Area aggregate and some specific national cases.

The Eurocoin approach is implemented in this thesis to estimate monthly smoothed GDP components for the following reasons:

1. Euro Area flash estimates are released on a quarterly frequency, with a certain delay (8-10 weeks);

2. GDP growth in any quarter depends on seasonal effects and measurement errors. On the contrary, Eurocoin methodology highlights the underlying trend by adjusting growth rate for short-term fluctuations and measurement errors.

3. Our disaggregated indicators can be influenced by factors affecting only a particular country or sector, differently from the classic Eurocoin indicator.

4. An alternative to Eurocoin in summarizing the current economic picture and in tracking underlying growth for Euro area could be the band-pass filters, but they can deteriorate at the end of the sample, and they are less reliable for the most recent data, and very relevant for economic policy. Differently, Eurocoin and the disaggregated indicators that we outline in this Thesis extract information relevant to the forecasting of GDP, preceding official releases and bandpass filters by several months.

This thesis contributes to the debate on the estimation of smoothed growth by generalized dynamic factor models.

We have analyzed our disaggregated indicators considering the following objects:

d) ability of the real time indicator to approximate the target as measured by the root mean-square error;

e) size of the revision errors after one month;

f) ability to signal the correct change of the bandpassed variation (see Peasaran and Timmermann, 1992);

g) detection of Turning Points (TP).

Comparing in chapter 3 real time estimates with bandpassed data (that have been analyzed in chapter 2 with the ex-post estimation), we observe, in particular, the following elements:

1) Estimating medium to long run component of the sectoral growth

- In terms of RMSFE and ability of real time indicator to track the direction of the target, the two aggregated macro indicators (concerning respectively Industry and Services) give better results in terms of RMSFE than the ones concerning specific sectors (Manufacturing, Construction, Energy, Financial). Perhaps, by aggregating data, volatility concerning sectoral cycles decreases as much as performance improves.
- We test the historical performance of Trade-Transport-Communication indicators and find that they do a good job at describing this sectoral business cycle. In the recursive real time estimation exercise, we find that the indicators generally produce a good sectoral GDP smoothed growth estimation.
- If a simple index concerning Trade-Transport-Communication appears particularly efficient to assess the underlying direction towards which these sectoral economies move, it appears more useful to assess other sectoral growth rates (Manufacturing, Energy, Construction) by the relative macro area indicators.
- Comparing pseudo real time estimates, we observe that the lower volatility of a bandpassed sectoral business cycle seems to improve the performance in real time of forecasting indicators (based on the generalized dynamic factor model).
- We also tested to eliminate production industrial data from our dataset, with the goal to build "Service" sector estimates. Results obtained show that by eliminating these data to build some forecasts concerning Financial, Trading, Transport and Communication sectors, does not change the structure of the common factors and does not improve RMSFE and correlation with sectoral bandpassed GDP.

2) National smoothed indicators:

- In terms of RMSFE and ability of real time indicator to track the direction of the target, Italy, Germany, France and Netherlands national indicators provide a satisfactory approximation of band-pass GDP in real time, in particular by the projection of National GDP on the 157 European variables included in the dataset (that produce what we call "European common factors").
- Interrelations between a large dataset of European variables and the National Gross Domestic Product seem to maximize the forecasting capacity of our indicators.
- The role of common factors in filtering data appears relevant; e.g., when we project a national GDP on European factors, we generally obtain, both by ex post estimation and in real time, an indicator more similar to Eurocoin.

- It seems possible to well estimate national smoothed growth by using a large dataset with 157 European variables. However, by using a more parsimonious number of variables (building some "National common factors", therefore only using variables in the Dataset regarding a specific country), it is also possible to obtain some consistent estimates (e.g. see Italian case in section 3.3).
- We also show that European variables (belonging to different countries) are essential to estimate national smoothed growths. In practice, we observe that underlying national common tendencies often go in the same direction and there are high interactions among European countries.

3) Combination of Real and Financial variables

- Pseudo real time estimation shows that, in the 2003-2008 period, a combination between "real" MLRG (medium to long run growth rate) and "financial MLRG", obtained projecting Euro Area GDP respectively on real and financial variables, can compete with the approximate ideal target within the sample, and it can slightly outperform Eurocoin in terms of RMSFE and correlation, particularly in 2005-2007, where band-pass monthly growth was particularly high.
- We can assess that the impact of real and financial variables in estimating smoothed GDP, during the structural break in 2008, shows that the role of real data as industrial production, demand indicators, foreign trade (Import, Export), and Employment Indexes, becomes particularly relevant in relation to that concerning financial data like Exchange rates, Money Supply, Spreads. Therefore, one possible explanation could be that interrelations among the recession phase and the variations in production, consumptions and unemployment are highly correlated.

4) Estimating smoothed Euro GDP components and National Consumptions

For Euro Area, both Final Consumptions and Household Consumptions have been particularly well forecasted in the recent past and we show ability of the real time indicator to *track the direction of the target.* Regarding Foreign Trade variables, the results of our estimates are not similarly clear and they must be investigated considering the high volatility of their business cycle. Regarding the capacity of each Euro Area GDP component on *nowcast of its bandpassed target*, the unprecedented recession of 2008-2009 strongly increases RMSE values. It also appears quite high for RMSE Investment.

Results concerning European aggregate show that it is not always clear whether it is more useful to use national or European variables to project Expenditure Components of GDP by the dynamic factor model. Thus, we have proposed to test different strategies.

In synthesis, it appears interesting to highlight that performances into the sample and in real time, with regard to both European growth rates and sectoral data, show how it is often useful to use a large dataset of European variables to estimate smoothed growth by the dynamic factor model. However, a more parsimonious and efficient use of variables is also a possibility to be considered in some specific cases. On the contrary, to estimate National Household Consumptions, the impact of Foreign (European) variables seems less relevant.

Clearly, such conclusions are not general but restricted to the data and the models used in these exercises.

Many extensions of the work developed in chapters 2 and 3 are possible. First of all, it could be interesting to analyze the impact of Euro Area, USA and UK macroeconomic variables to estimate the medium to long run component of the growth for Eastern European EU member countries (i.e. Poland and Slovenia), also to analyze interactions among different business cycles.

Furthermore, a possible extension could be to study the impact of the “Spread variable” to forecast smoothed growth by generalized dynamic factor model. Since the 1980s, econometric literature has developed in support that the slope of the yield curve, i.e. the spread between long and short-term interest rates, is a good predictor of recessions and future economic activities.

An other interesting research question would be: how real and financial variables that we have used in estimating smoothed GDP could be useful to analyze inflation rate by generalized dynamic factor model?

Finally, an interdisciplinary remark: Euro Area does not have a strong Economic Government, however econometric research shows how it is already a very interrelated system of countries.

BIBLIOGRAPHY

Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, L. Reichlin, and G. Veronese (2001). A real Time Coincident Indicator for the euro area Business Cycle. CEPR Discussion Paper No. 3108.

Altissimo, F., R. Cristadoro, M. Forni, M. Lippi, and G. Veronese (2010). New Eurocoin: Tracking Economic Growth in Real Time. *The Review of Economics and Statistics, MIT Press, vol. 92(4)*, pages 1024-1034.

Amengual, D. and M. Watson (2007). Consistent estimation of the number of factors in a large N and T panel". *Journal of Business & Economic Statistics, 25*, 91-96.

Anderson, T. W. (1984). An Introduction to Multivariate Statistical Analysis. New York: John Wiley & Sons.

Angelini, E., J. Henry and M. Marcellino (2006). Interpolation and backdating with a large information set. *Journal of Economic Dynamics and Control*, *Vol. 30, No. 12*, pp. 2693-2724.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Runstler (2008). Short-run forecasts of euro area GDP growth. CEPR discussion paper series n. 6746.

Aruoba, S.B., F.X. Diebold, and C. Scotti (2009). Real-Time Measurement of Business Conditions. *Journal of Business and Economic Statistics, 27*, 417-427.

Baffigi, A., R. Golinelli, and G. Parigi (2004). Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, *Vol. 20*, pp. 447–460.

Baffigi, A., R. Golinelli, and G. Parigi (2002). Real-time GDP forecasting in the euro area. Economic working papers 456, Bank of Italy, Economic Research Department.

Bacchini,F., A. Ciammola, R. Iannaccone and M. Marini (2010). Combining forecasts for a flash estimate of Euro area GDP. Contributi (WP) Italian National Statistical Institute (ISTAT) n.3/2010.

Backus D.K. and P.J. Kehoe (1992). International Evidence on the Historical Properties of Business Cycles. *The American Economic Review*, 82*,* 864-888*.*

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*, 191-221.

Bai, J. (2003). Inferential theory for factor models of large dimension. *Econometrica 71*, 135-71.

Bai, J. and S. Ng (2005). Tests for Skewness, Kurtosis, and Normality for Time Series Data. *Journal of Business & Economic Statistics*, *American Statistical Association, vol. 23*, pages 49-60

Bai, J. and S. Ng (2007). Determining the Number of primitive Shocks in Factor Models. *Journal of Business and Economic Statistics, 25(1), American Statistical Association*, pp.52-60.

Banbura, M., D. Giannone, L. Reichlin (2010). Nowcasting. CEPR Discussion Papers 7883.

Bank of Italy Official Site. *New Eurocoin, A tutorial Note*. Available at http://eurocoin.bancaditalia.it

Barhoumi, K., S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, G. Runstler, K. Ruth, and C.Van Nieuwenhuyze (2008). Short-term forecasting of GDP using large monthly datasets: a pseudo real-time forecast evaluation exercise. European Central Bank Occasional Paper, No. 84.

Baxter, A. and R.G. King (1999). Measuring Business Cycle Approximate Band-Pass filters for Economic Time Series. *The Review of Economics and Statistics 81*,575-9343

Bernanke, B. and J. Boivin (2003). Monetary Policy in a Data-Rich Environment. *Journal of Monetary Economics* 50:3, 525-46.

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 Measurement of the Business Cycle. *Journal of Monetary Economics*, 7(2), 151–74.

Billmeier, A. (2004). Ghostbusting: Which Output Gap Measure Really Matters?. IMF Working Papers, n.146.

Boragan, S. Aruoba, and Francis X. Diebold 2010. Real-Time Macroeconomic Monitoring: Real Activity, Inflation, and Interactions. *American Economic Review*, Vol. 100, No. 2 (May), pp. 20-24.

Brillinger, D. (1981). Time series – Data Analysis and Theory, expanded ed., San Francisco.

Bry, G., and C. Boschan (1971). Cyclical Analysis of Time Series: Selected Procedures and Computer Programs. National Bureau of Economic Research, New York.

Brockwell, P.J. and R. A. Davis (1991). Time Series: Theory and Methods. 2nd Edition. New York: Springer-Verlag.

Bruno, G., Lupi, C. (2002). Forecasting Industrial Production and the Early Detection of Turning Points. Università del Molise, Dipartimento di Scienze Economiche Gestionali e Sociali, Quaderni di Ricerca, Serie Statistica, n. 4.

Buffettau, S. and Mora, V. (2000). Predicting the national accounts of the Euro zone using business surveys. Conjoncture in France, INSEE.

Burns, A.F., and W.C. Mitchell (1946). Measuring Business Cycles. New York, NBER.

Camacho, Maximo, and Gabriel Perez-Quiros (2010). Introducing the euro-sting: Short-term indicator of euro area growth. *Journal of Applied Econometrics*, *Vol. 25, No. 4*, pp. 663-694.

Canova, F., Paustian, M. (2007). Business cycle measurement with some theory. Economics Working Papers 1203, Department of Economics and Business, Universitat Pompeu Fabra, revised Jul 2011.

Chadha, S., N. Janssen, C. Nolan (2000). An examination of UK business cycle fluctuations: 1871-1997. Discussion paper in economics and management, University of Reading.

Chamberlain, G. (1983) Funds, factors, and diversification in arbitrage pricing models. *Econometrica* 51, 1281-1304.

Chamberlain, G., and Rothschild, M. (1983). Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets. *Econometrica*, 70, 191–221.

Clements M.P. and Hendry, D. F.(2004). Pooling of Forecasts. *Econometrics Journal, 7, 1*, 1-31.

Christiano, L.J. and Fitzgerald, T.J. (2003). The Band-Pass Filter. *International Economic Review* 84, 435-65.

Cristadoro, R., M. Forni, L. Reichlin, and G.Veronese (2005). A core inflation indicator for the euro area. *Journal of Money Credit and Banking 37(3)*, 539-560.

Cristadoro, R., Secchi, A., Veronese, G. (2008). L'indice punta alla recessione. Available at www.lavoce.info

Diebold, F. and R.S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics 13*, 253-263.

Diebold, F. and G. Rudebusch (1996). Measuring Business Cycles: a Modern Perspective. *The Review of Economics and Statistics, 78*, pp. 66-77.

Diron M. (2006): Short-term forecasts of euro area real GDP growth. An assessment of real-time performance based on vintage data, ECB. Working Paper series No. 622.

Dolado, J.J., M. Sebastián, J. Vallés (1993). Cyclical Patterns of the Spanish Economy. *Investigaciones Económicas*, *XVII,* 445-473.

Doz C., D. Giannone and L. Reichlin (2006). A quasi maximum likelihood approach for large approximate dynamic factor models. CEPR Discussion Paper 5724

Doz, C., D. Giannone, and L. Reichlin, (2007). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. Discussion paper 6043, Centre for Economic Policy Research (CEPR).

European Central Bank (ECB) Monthly Bulletin – April 2008

Fagan, G., J. Henry, and R. Mestre (2001). An area-wide-model (AWM) for the euro area. ECB Working Paper Series no. 42.

Fagan G. e J. Henry (1998). Long run money demand in the EU: evidence from area-wide aggregates. *Empirical Economics*, Vol. 23**,** pp. 483-506.

Fase M. M. G. e C. C. A. Winder (1998), "Wealth and the demand for money in the European Union", *Empirical Economics*, Vol. 23, pp. 507-524.

Favero, C., Marcellino, M., Neglia, F.(2005). Principal components at work: the empirical analysis of monetary policy with large datasets. *Journal of applied Econometrics*, pp.603- 620, Vol.20.

Forni, M. and L. Reichlin (1996). Dynamic Common Factors in Large Cross-Sections. *Empirical Economics, 21*, 27-42.

Forni, M. and L. Reichlin (1997). National Policies and Local Economies: Europe and the United States. CEPR Discussion paper no.1632.

Forni M. and Lippi M. (1997). *Aggregation and the Microfoundations of Dynamic Macroeconomics*. Oxford: Clarendon Press.

Forni, M. and L. Reichlin (1998). Lets Get Real: A Dynamic Factor Analytical Approach to Disaggregated Business Cycle. *Review of Economic Studies, 65,* 453-474.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalized dynamic factor model: identification and estimation. *The Review of Economics and Statistics 82*, 540-554.

Forni M., M. Hallin, M. Lippi and L. Reichlin (2000). Reference Cycles: The NBER Methodology Revisited. CEPR Discussion Paper 2400.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2001). Coincident and leading indicators for the euro area. *The Economic Journal 111*, 62-85.

Forni, M. and M. Lippi (2001). The generalized dynamic factor model: representation theory. *Econometric Theory 17*, 1113-41.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2004). The generalized dynamic factor model: consistency and rates. *Journal of Econometrics 119*, 231-255.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005). The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association 100*, 830-40.

Fourier, J., (1822). The Analytical Theory of Heat, English translation by Freeman (1878), republication by Dover, New York (1955).

Fournier, J.Y. (1999). Extraction du cycle des affaires: la méthode de Baxter e King. *Série des documents de travail de la Direction des Etudes et Synthèses Économiques*, *INSEE.*

Frenda, A. (2011). Analyzing National smoothed growth by generalized dynamic factor model. *International Journal on GSTF Business review* , Vol.1, No.1, August 2011

Frenda, A. (2012). Estimating smoothed GDP components*. China-USA Business Review*, David Publishing Company, Vol. 11, No.1

Fukuda, K. (2009). Forecasting growth cycle turning points using US and Japanese professional forecasters. *Empir Econ 36*:243–267

Giannone, D., Lenza, M. (2004). The Feldstein-Horioka Fact. CEPR Discussion Papers 4610

Giannone, D., L. Reichlin, L. Sala (2005). Monetary Policy in Real Time. CEPR Discussion Papers 4981

Giannone, D., L. Reichlin, & L. Sala (2006). VARs, common factors and the empirical validation of equilibrium business cycle models. *Journal of Econometrics* 127, 257–279.

Giannone, D. and L. Reichlin (2006). "Does information help recovering structural shocks from past observations?". *Journal of the European Economic Association*, 4(2-3), 455-465.

Giannone D., L. Reichlin and D. Small (2006). Nowcasting GDP and inflation. The real-time informational content of macroeconomic data releases. ECB Working Paper series No. 633, May 2006.

Giannone, D. and T. Matheson (2007). A New Core Inflation Indicator for New Zealand. *International Journal of Central Banking, 37(3)*, pp. 145-180.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, Vol. 55, No. 4, pp. 665 – 676.

Giannone, D., M. Lenza and L. Reichlin (2008). Explaining the great moderation: it is not the shocks. *Journal of the European Economic Association, P&P, Vol. 6, No. 2-3*, Pages 621-633.

Giannone D., M. Lenza and L. Reichlin (2009). Business Cycles in the Euro Area. WP Series No 1010.

Giorno, Cl., Richardson, P., Roseveare, D. and Van Den Nooord, P. (1995). Estimating Potential Output, Output Gaps and Structural Budget Balances. OECD Economics Department Working Papers, 152.

Geweke J. (1977). The Dynamic Factor Analysis of Economic Time Series Models, in Latent Variables in Socioeconomic Models, North Holland, pp.365-383.

Granger, C. W. J., and O. Morgenstern (1963). Spectral Analysis of New York Stock Market Prices. *Kyklos*, 16(1), 1-27.

Granger, Clive W.J. and M. Hatanaka (1964). *Spectral Analysis of Economic Time Series* (Princeton, NJ: Princeton University Press).

Granger, CWJ (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica, 50*, pp. 424-38.

Gregoir, S. and G. Laroque (1992). La place des stocks dans les fluctuations conjoncturelles. *Annales d'économie et de statistique,* n. 28

Grenouilleau D. (2004). A Sorted Leading Indicators Dynamic (SLID) Factor Model for short-run Euro-area GDP Forecasting. European Commission Economic Paper 219.

Hallin, M. and R. Liska (2007). Determining the Number of Factors in the Generalized Factor Model. *Journal of the American Statistical Association, 102, No. 478*, 603-617.

Harding, Don & Pagan, Adrian (2002). Dissecting the cycle: a methodological investigation. Journal of Monetary Economics. *Elsevier, vol. 49(2)*, pages 365-381

Harvey, A.C. (1981). Time Series Models (Philip Allan Publishes Limited, Oxford)

Henderson, R. (1916). Note on Graduation by Adjusted Average. *Transactions of the American Society of Actuaries, 17*, 43-48. Note: a reproduction of this paper is available in Appendix 2 of Australian Bureau of Statistics (2003).

Hodrick RJ and EC Prescott (1997). Post-war US business cycles: an empirical investigation. *J Money Credit Bank 29*:1–16

Lewbel, Arthur (1991). Cost of characteristics indices and household equivalence scales. *European Economic Review, Elsevier, vol. 35(6)*, pages 1277-1293

Lucas, Robert E. (1977). Understanding business cycles. Carnegie-Rochester Conference Series on Public Policy, *Elsevier, vol. 5(1)*, pages 7-29, January.

A.Maravall, A. del Rio (2001). Time Aggregation and the Hodrick-Prescott Filter. *Banco de Espana Working Paper Series* 2001 (08)

Mariano, Roberto S., and Yasutomo Murasawa (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, *Vol. 18, No. 4,* pp. 427-443.

Kalman, R.E.(1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions ASME Journal of Basic Engineering 82*, 35-45.

Kapetanios, G., M. Marcellino (2004). A parametric estimation method for dynamic factor models of large dimensions. Queen Mary University of London Working Paper 489.

Kapetanios G. (2004). A new method for determining the number of factors in factor models with large datasets. Working paper No 525, Department of Economics, Queen Mary University of London.

Kapetanios G. (2005). A testing procedure for determining the number of factors in approximate factor models with large datasets. Working paper No 551, Department of Economics, Queen Mary University of London.

Kitchen, J., Monaco, R. (2003). Real-Time Forecasting in Practice: The U.S. Treasury Staff's Real-Time GDP Forecast System. MPRA Paper 21068, University Library of Munich, Germany, revised Oct 2003

Matheson, Troy D. (2010). An analysis of the informational content of New Zealand data releases: The importance of business opinion surveys. *Economic Modelling*, Vol. 27, pp. 304-314.

Matheson, T. 2011. New Indicators for Tracking Growth in Real Time. IMF Working Papers 11/43, International Monetary Fund.

Mitchell, W. C. (1927): Business Cycles: The Problem and Its Setting. National Bureau of Economic Research, New York.

Mitchell, J., Smith R., Weale M., Wright S. and Salazar E. (2004). An Indicator of Monthly GDP and an Early Estimate of Quarterly GDP Growth. National Institute of Economic and Social Research Discussion Paper.

Monticelli C., M.O. Strass-Kahn (1992). European Integration and the Demand for Broad Money. *BIS, Working Paper*, n. 19.

Nguiffo, Bojom (2008). A monthly indicator for economic activity for Luxembourg. Banque Centrale du Luxembourg. Working Paper No. 31.

OECD, Monthly Note of Economic Leading Indicators.

Onatski A. (2006a). Asymptotic distribution of the principal components estimator of large factor models when the factors are relatively weak. Mimeo

Onatski A. (2006b). Determining the number of factors from empirical distribution of eigenvalues. Mimeo

Onatski A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica, 77,* 1447-1479.

Pesaran, M.H. and A. Timmermann (2006). Testing dependence among serially correlated multi-category variables. CESifo Working Papers, 1770.

Quah, D. (1992). The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds. *Econometrica* 60, 107-118.

Quah, D. & Thomas J. Sargent (1993). A Dynamic Index Model for Large Cross Sections. CEP Discussion Papers dp0132, Centre for Economic Performance, LSE.

Reichlin, L. (2007). *The Eurocoin project: new modelling ideas at central banks* - CEPR

Rünstler G., F. Sédillot (2003). Short-term estimates of euro area real GDP by means of monthly data. Working Paper Series 276, European Central Bank.

Sargent, Thomas J. (1973). Interest Rates and Prices in the Long Run: A Study of the Gibson Paradox. *Journal of Money, Credit and Banking,* Blackwell Publishing, vol. 5(1), pages 385-449

Sargent, T. J. and C. A. Sims (1977). Business Cycle Modeling Without Pretending to have too Much a priori Economic Theory. *New Methods in Business Cycles Research*, Federal Reserve Bank of Minneapolis, pp. 45-110.

Sargent, Thomas J. (1987). *Macroeconomic Theory*, Second Edition. Academic Press, San Diego, CA.

Stock, J. H., and Watson, M. W. (1989). New Indexes of Coincident and Leading Economic Indicators. *NBER Macroeconomics Annual*, pp. 351-393.

Stock, J. H., and M. W. Watson (1998). Business Cycle Fluctuations in U.S. Macroeconomic Time Series. Discussion Paper 6528, National Bureau of Economic Research, Inc.

Stock, J. H., and Watson, M. W. (1999). Business Cycle Fluctuations in U.S. Macroeconomic Time Series. In G. B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, pp. 3-64. Amsterdam: Elsevier Science Publishers.

Stock, J.H. and M.W. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*, 1167-79.

Stock, J.H. and M.W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics 20*, 147-162.

Valle e Azevedo J., Koopman S.J., Rua A. (2006). Tracking the business cycle of the euro area: a multivariate model-based band-pass filter. *Journal of Business and Economic Statistics, 24*, 278-290.

Zarnowitz V (1969). The new ASA-NBER survey of forecasts by economic statisticians. *Am Stat 23*: 12–16

Zarnowitz, V. and G. H. Moore. (1986), Sequential signals of recession and recovery. *Journal of Business, Vol. 55*, 57-85.

Zarnowitz, V. (1992). *Business Cycles: Theory, History, Indicators, and Forecasting*. The University of Chicago Press, Chicago.

Zarnowitz, V. and Ozyildirim, A. (2002). Time series decomposition and measurement of business cycles, trends and growth cycles. NBER WP No. 8736.

Zarnowitz V, Ozyildirim A (2006). Time series decomposition and measurement of business cycles, trends and growth cycles. *J Monet Econ 53*:1717–1739