RESEARCH IN OFFICIAL STATISTICS

2004 EDITION

Monographs of official statistics

Papers and proceedings of the third Eurostat colloquium on modern tools for business cycle analysis

COPYRIGHT
Luxembourg: Office for Official Publications of the European Communities, 2004

ISBN 92-894-7805-5
ISSN 1725-5406
Cat. No. KS-CR-04-001-EN-N

1

EUROPEAN COMMISSION

eurostat

THEME 1
General statistics

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

New freephone number:

00 800 6 7 8 9 10 11

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

Luxembourg: Office for Official Publications of the European Communities, 2004

ISBN 92-894-7805-5
ISSN 1725-5406

Monographs of official statistics

Papers and proceedings of the Third Eurostat Colloquium on Modern Tools for Business Cycle Analysis

Statistical methods and business cycle analysis of the Euro zone

Edited by

Gian Luigi Mazzi and Giovanni Savio

Acknowledgements

This publication contains the proceedings of the third Colloquium on Modern Tools for Business Cycle Analysis held in Luxembourg, 28-29 November 2002.

Project management and coordination were ensured by Klaus Reeh.

Eurostat would like to thank all the actors involved in the writing and preparation of this publication, in particular the authors of the papers, and Nelly Da Silva and Jane Schofield for the support and the organisation provided.

Contents

Foreword

Why a Eurostat Colloquium on "Modern Tools for Business Cycle Analysis"

By Klaus Reeh

Adviser, Eurostat, Luxembourg

Econometric tools for business cycle analysis are used, of course, by economists when they work on the series provided by official statisticians. However, they are also used by official statisticians themselves and this basically for three purposes.

First they subject their series to standard transformation procedures such as seasonal adjustment, outlier detection or detrending to make them more easily readable. We at Eurostat consider this to be a service to be rendered for our EU/EMU series. Such a service can go as far as suggesting a business cycle dating based on statistical techniques. We have the impression that an upgrade of this kind of service is needed by going from simple standard transformation procedures to somewhat more sophisticated procedures.

Second and quite closely related to the first reason, they have to check the quality of their series. Tools for analysing series are, of course, also helpful when it comes to checking them. As it stands we at Eurostat are not yet as developed in our quality checking, as we would have liked to be, most notably in business cycle statistics. We hope, however, to profit in our work from the analytical tools either already existing or under development.

Third, and this is a special problem for official statisticians dealing with EU/EMU series, our series have some shortcomings that could be remedied at least partially with the help of analytical tools. We have to overcome the lack of stability in time and thus insufficient series length affecting most of our series with retrapolation tools that are, of course, in their structure quite close to extrapolation tools. Similarly we have to handle interpolation problems. There are cases where not all Member States provide data in the desired frequency (e.g. only quarterly instead of monthly, only annual instead of quarterly series). Here again the analytical tools turn out to be helpful.

Finally our series suffer quite often from a lack of timeliness. We have sometimes data not from all Member States. Therefore we have to be inventive and come up with flash estimates (not forecasts, but coincident or leading indicators). Here again the tools to be used are the very same that can be used for forecasting and extrapolation.

Official statisticians at Eurostat have therefore a strong interest in what kind of tools are currently discussed, developed, tested and used. Closer ties with the scientific community in general and the business cycle analysis community in particular are very helpful, as Eurostat entered only recently into the business of business cycle statistics with its Eurozone indicators. Let us therefore hope for more such colloquia in the future, as both the scientific community and official statisticians can only profit from it.

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002

STATISTICAL METHODS FOR BUSINESS CYCLE ANALYSIS

By Gian Luigi Mazzi and Giovanni Savio

Statistical Office of the European Communities, Eurostat[1]

1 Introduction

For a long time period business cycle analysis has been considered the most empirical part of economics. In effect, being free of any specific economic theory has been the leitmotiv of the researches conducted in this field since the advent of the economic barometers developed by Persons (1919) within the Harvard Committee of Economic Research at Harvard University. The atheoretical philosophy has been successively strenghtened in the pioneering works of Burns, Mitchell and their colleagues at the NBER. The empirical characterization of the inductivist approach has also influenced the way in which the economic analyses were carried out for some decades, as only elementary statistical techniques were used, and a central role was played by individual judgment.

The situation gradually started changing from the end of the '70s onwards. On one side, real business cycle theories have created new links between economic theory and business cycle analyses. On the other side, various studies have successfully formalized in a statistical framework the approaches commonly used for business cycle studies. In the meantime, statisticians and econometricians have improved business cycle analyses by either introducing more sophisticated techniques (Stock and Watson (1991), Hamilton (1989)), or by using less traditional definitions of key concepts, such as trends and cycles (Beveridge and Nelson (1981), Vahid and Engle (1993)).

Thanks to all these changes several important improvements have been achieved so far, amongst them: a stochastic generalization of key concepts such as trends, cycles and turning points; a more robust estimation framework of classically defined trends and cycles due to the use of sophisticated filtering techniques; a well defined set of algorithms - both parametric and non-parametric - to date, detect and forecast turning points.

[1] Unit C6, Statistical indicators for the Euro zone. Address: Bech Building, 5, rue A. Weicker, L-2920 Luxembourg. E-mail: gianluigi.mazzi@cec.eu.int and giovanni.savio@cec.eu.int.

In this pages we propose a general overview of such improvements by giving key-readings and syntheses of the papers presented at the Colloquium on Modern Tools for Business Cycle Analysis, organized by Eurostat in Luxemburg on November 2002. As such, our introduction does not claim to be exhaustive at all. It touches specific fields - those discussed during the Colloquium - located at the frontier of business cycle analysis and on which the theoretical and empirical debate is still open.

The scheme of this paper closely follows the contents of the sessions organized during the Colloquium. Section 2 analyses developments in detrending techniques. Section 3 presents an assessment of recent researches in turning points dating and detection. Section 4 discusses contributions on convergence and synchronization of economic movements. Section 5 deals with the use of multivariate techniques. Section 6 briefly concludes.

2 Detrending techniques

Growth or deviation cycles (and the corresponding concept of output gap) are often considered as key variables for economists and policy makers, especially in the definition of monetary policies aimed at controlling the inflation rate. Unfortunately, the sensibility of growth cycle estimates to detrending techniques and their instability for current periods can considerably reduce their relevance for practical purposes. For example, the recent literature has focused the attention on the causes of such instabilities as well as on their relative importance in explaining biases in preliminary estimates (see, e.g., Orphanides and van Norden (2002)). In the same respect, the results of many studies have shown that multivariate decomposition approaches can outperform univariate models in terms of reliability, revision properties, and inflation forecasts. Multivariate decomposition approaches have as their main rationale that adding economic content to the estimations can improve the trend-cycle decomposition through a better disentanglement of supply and demand shocks. Then, it has become quite common to encounter extensions of the literature allowing for the inclusion in estimated models of a Phillips' curve, the Nairu or the Okun's law.

The paper "The impact of the macroeconomic hypothesis on the estimation of the output gap using a multivariate Hodrick-Prescott filter: the case of the Euro area" by Odile Chagny and Mathieu Lemoine analyses - through state-space techniques - the impacts of integrating alternative macroeconomic hypotheses on the estimation of the output gap of the Euro zone using some multivariate extensions of the Hodrick-Prescott's filter (see Laxton and Tetlow (1992)). In line with previous researches, the authors find that integrating macroeconomic information can in some specific periods significantly modify the appreciation of the output gap level of the Euro zone given by the univariate Hodrick-Prescott filter. Furthermore, the assessment of both the reliability of the alternative output gaps and of the revision properties shows a substantial superiority of models based on the capacity utilization rate over the other models.

The use of dynamic factor analysis and dynamic principal component techniques to synthesize large data-sets and to extract common elements from them is very well-known and developed in econometrics. Nevertheless, many of these methods - especially those using state space model estimated with maximum likelihood - become soon not practical when the dimension of the model increases too much. In his paper "Modelling core inflation for the UK using a new dynamic factor estimation method and a large disaggregated price index dataset" George Kapetanios discusses an alternative method for estimating factors derived from a factor state-space model. His model has the following advantages: it has a clear dynamic interpretation, it does not require iterative estimation techniques, and it can accommodate cases where the number of variables exceeds the number of observations.. Further, as the factor analysis is carried out by the author within a general model, forecasting is easier to carry out than in the currently available procedures, where a forecasting model needs to be specified. Kapetanios, after discussion of the merits of the proposed approach, presents an application of the suggested method to the extraction of core inflation and forecasting of UK inflation in the recent past. The measure of core inflation obtained is shown to have predictive ability for inflation in the UK over a relatively long evaluation period.

Torben Mark Pedersen, in his paper "Alternative linear and non-linear detrending techniques: a comparative analysis based on Euro-zone data", compares the distortionary effect of various filters proposed by the literature using quarterly real GDP for the EU, the Euro-zone, and a number of EU countries. His measurements are based on the *ad hoc* assumption that the 'true' business cycle filter is an ideal high-pass filter, or an ideal band-pass filter for a band of frequencies which is determined after measuring the duration of growth cycles or the business cycle component. Unfortunately, the author does not find general results about the distortionary effect of the examined filters, but he shows that biases are in many cases substantial, that these depend on the time series being filtered, and that different filters may be optimal for different countries or different time series.

3 Turning points detection

The identification of turning points is at the the centre of almost all business cycle analyses. Since the publication of the pioneering book of Burns and Mitchell (1946), a particular attention has been paid by business cycle researchers to the definition of turning points, to their location in the past (dating of turning points), to their real time identification (detection), as well as to their anticipation (use of leading cyclical indicators). Though the classical definition of cycles and the assessment of turning points proposed by Burns and Mitchell (1946) still remain the most commonly used, alternative definitions of turning points based either on growth cycles, on acceleration cycles, or on the so-called recovery cycles, have been developed by the recent literature. In particular, due to the length of cycles observed during the last years, there is an increasing attention of economists and policy makers to turning points observed on the growth cycles, namely on the shifts between acceleration and deceleration phases.

Concerning the methodology, two main approaches have been developed. The first is based on non-parametric techniques, and has its roots in the algorithm developed by Bry and Boschan (1971). The second is mainly founded on non-linear parametric techniques, such as switching-Markov models, or threshold models. These approaches try to replicate the results obtained by the NBER dating committee and other dating committees around the world. Usually, classical and growth cycle approaches are considered as distinct alternatives for business cycle analysis.

In their paper entitled "A comparative assessment of parametric and non-parametric turning points detection methods: the case of the Euro-zone economy", Jacques Anas and Laurent Ferrara present their integrated approach to turning points identification based simultaneously on classical, growth and acceleration cycles. A theoretical sequence of turning points based on 3 cycles is presented and a clear interpretation is given. Moreover, authors provide a general assessment of both parametric and non-parametric methods, and give practical recommendations on which of the two approaches is likely to give better results in dating and detecting contexts.

The main features and characteristics of non-parametric turning points methods are presented in the paper by Donald Harding "Non-parametric turning point detection, dating rules and the construction of the Euro-zone chronology", with an interesting comparison with parametric ones, where - following Harding - it is common to equate decisive changes with regime changes as in the case of the Markov-switching (MS) models. What constitutes a decisive change in the series, therefore a turning point, is the key issue to be resolved prior to formalizing and implementing a definition of turning points. Harding starts in presenting three markers that are based on successively weaker notions of what constitutes a 'decisive change', and formalizes them into non-parametric procedures for locating turning points. Then, he applies these procedures to construct GDP-based business cycle chronologies for several European economies and the Euro zone as a whole. Once turning points have been located in several series, it is of interest to investigate whether there exists a common cycle and to study the synchronization between the common cycle and the specific cycles from which it was constructed. A method for aggregating turning points is discussed and then applied to European data to obtain an Euro zone reference cycle. The common cycle located by aggregating turning points is compared with the common cycle located in the aggregate of Euro zone GDP.

Opposite theoretical and empirical results are found by Hans Martin Krolzig in the paper entitled "Constructing turning point chronologies with Markov–switching vector autoregressive models: the Euro–zone business cycle". The proposed approach is based on the MS time series model originally formulated by Hamilton (1989) and generalized to a MS vector autoregressive model, which is used for the modelling and dating of the Euro zone business cycle. The parametric approach for the constructing the turning point chronology consists of the following phases: (i) modelling the Euro zone business cycle as a single common factor generated by a hidden Markov chain; (ii) fitting a congruent statistical model to the data, (iii) deriving the conditional probabilities of the regimes; (iv) classifying each point in time to the regime with the highest probability; (v) dating the turning points of the business cycle. The MS-VAR also provides measures of uncertainty associated with the turning point chronology, facilitates real-time detection of business cycle transitions, and offers a well-developed theory for the prediction of the business cycle. In the empirical part, the MS-VAR approach is applied to three multi-country data sets consisting of real GDP and industrial production growth rates for the Euro zone.

4 Cyclical convergence and forecasting

Convergence and synchronization are at the core of Burns and Mitchell (1946, pg. 3) definition of business cycles:

> Business cycles are a type of fluctuations found in the aggregate economic activity of nations . . . : 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.

Then, a clear relationship in terms of synchronization among different economic series is required. Moreover, business cycles are defined as fluctuations of a latent series synthesizing the evolution of many series: individual cycles observed in many series should converge to a common unobserved cycle.

These concepts have been developed during the last twenty years in several methodological frameworks, particularly: geographical convergence (Barro and Sala-i-Martin (1992)); cointegration (Engle and Granger (1987)); common features and cycles (Vahid and Engle (1993)); static/dynamic factor analysis (Stock and Watson (1991)); structural time series models (Harvey (1989)). In empirical works, a number of papers have used these techniques in order to test for convergence and to investigate on the existence of commonalities among trends and cycles of EU countries. Synchronization/convergence analyses will certainly continue in being a key issue after future enlargements of the EU.

In their paper "Convergence and cycles in the Euro-zone", Vasco Carvalho and Andrew Harvey establish stylized facts about convergence in Euro zone countries both with respect to long-run income levels and to cycles. Their analysis is based on a new multivariate unobserved components model in which convergence components, formulated through a second-order error correction mechanism able to capture temporary disequilibrium, are combined with a common trend and similar cycles. The results obtained for eleven Euro zone countries indicate the existence of two convergence clubs, a high income group and a low income group. The multivariate convergence model is successful in separating trends from cycles and capturing the absolute convergence in the two groups. The evidence for convergence in the low income group is less compelling, but the assumption of a single common trend is found to be not unreasonable. The groups themselves appear to have converged in the relative sense. The implication is that the average per capita income in the poor group will remain almost 30% below that of the high group. The cycles in the core high income group show a remarkable coherence in recent years, with the group standard deviation having fallen dramatically. There is less coherence in the poor group, though again there is evidence of a movement towards the same cycle as the rich group in recent years.

Synchronization and convergence can represent a very helpful element to be included in forecasting models. The paper by Massimiliano Marcellino "Forecasting monthly macroeconomic variables for the Euro area" compares the linear, time-varying, and non-linear forecasts for aggregate monthly Euro zone macroeconomic variables obtained by two alternative aggregation procedures. The first uses constant aggregation weights, while the second is founded on time-varying weights. The main result is that for several variables forecasts from linear models can be substantially improved upon by using nonlinear or time-varying models, a circumstance of some interest for policy purposes but also - more generally - for empirical macroeconomic analyses. For example, it suggests that measures of persistence based on linear specifications can be inappropriate, as well as impulse response functions.

An interesting analysis on the existence of a common European cycle and on the differences induced by alternative detrending techniques is contained in the paper "Is there a common Euro-zone business cycle?" by James Mitchell and Kostas Mouratidis. The authors, using 40 years of monthly industrial production indices data and alternative parametric and nonparametric univariate measures of the 'classical' and 'growth' cycles, examine the relationship between the business cycles of the 12 Euro zone countries and investigate both whether their cycles have moved 'closer' together, and whether they have become more correlated over time. Authors' findings are that business cycles display different properties and that the degree of correlation varies according to the measure considered. However, their measures of synchronization between Euro zone business cycles, namely the average (size-weighted) correlation between them, exhibits common features across alternative measures of the business cycle. This increased synchronization between Euro-zone business cycles is consistent with the emergence of a 'common' Euro zone business cycle. Accompanying this increased correlation, their results indicate that Euro-zone business cycles have moved 'closer' together.

5 Multivariate decomposition methods

The simultaneous modelling of growth and cycle for a large set of series represents a very useful contribution in the context of business cycle analyses. It gives the opportunity of taking into account all the likely interactions among variables and gives a more sophisticated and complete view of the economic behavior. Relevant contributions in this field have been produced, amongst others, by Laxton and Tetlow (1992) with the multivariate generalization of the Beveridge and Nelson (1981) decomposition, by Evans and Reichlin (1994) with the multivariate Hodrick and Prescott (1997) filter, by Harvey (1989) with the multivariate extension of the structural time series approach, by Blanchard and Quah (1989) with the application of structural VAR models to business cycle analysis, by Vahid and Engle (1993) in the common trends-common cycles multivariate decomposition, etc..

A general multivariate model describing the Euro zone behavior is presented by Siem Jan Koopman in the paper "The common converging trend-cycle model: estimation, modelling and an application". In his model, Koopman includes and test for several constraints among variables, such as convergence of their long- and short-run movements, as well as their synchronization. He discusses multivariate time series models based on unobserved components with dynamic converging properties. Koopman defines convergence in terms of a decrease in dispersion over time, and models this decrease through mechanisms that allow for gradual reductions in the ranks of covariance matrices associated with the disturbance vectors driving the unobserved components of the model. The inclusion of such convergence mechanisms makes the identification of the various types of convergence possible. The multivariate unobserved common converging component model is applied to the per-capita GDP for five European countries: Germany, France, Italy, Spain and the Netherlands.

In the paper "State space decomposition under the hypothesis of non zero correlation between trend and cycle with an application to the Euro-zone", Tommaso Proietti discusses several issues related to trend-cycle decompositions with correlated components of macroeconomic time series, and illustrates them with reference to the Euro zone and the Italian GDP. In particular, he addresses the small sample properties of the estimated correlation of the trend and cycle disturbances, and reviews the interpretative issues raised by these decomposition. The nature of inferences about trends and cycles, with reference to the real time and final estimates, and the related topic of revision, is considered, along with the relationship with other popular results, such as the Beveridge and Nelson (1981) decomposition, the Single Source of Error and the Innovation models. He also looks at the consequences of seasonal adjustment and temporal aggregation on the empirical evidence for a negative correlation between the disturbances. Finally, he illustrates that multivariate analysis can provide additional insight on these topics.

6 Conclusions

Sophisticated statistical and econometric methods are nowadays available for business cycle analysts. These methods can integrate and improve the appreciation of economists on the current state and future development of the economy. In this context, a more active role of official statisticians can be sketched out. Official statisticians have traditionally been requested to do their best to fulfil economists' needs. In particular, they are asked to provide long and timely time series in order to achieve both a historical and a real time analysis of economic movements. Moreover, they are also required to provide transparent and well-documented seasonally adjusted data which are traditionally at the core of business cycle analyses. Furthermore, thanks to the development of new statistical and econometric techniques, official statisticians can play a major role in extending scope and coverage of data, and to propose alternative estimation strategies. For example, they can provide useful inputs to economists for their analyses, such as reliable and real time estimates of the output gap, turning points, convergence and synchronization measures.

Then, contributions from official statisticians will be interpreted and validated by economists in order to produce a reliable assessment of the cyclical situation of the economy. A solid cooperation between official statisticians and economists will probably lead to a more robust framework for business cycle analyses. The expected result of such cooperation could be a more reliable and transparent set of indicators to be supplied to policy and decision makers in order to help them in the definition of current policy strategies.

References

BARRO, R. J., AND X. SALA-I-MARTIN (1992): "Convergence," *Journal of Political Economy*, 100, 223-251.

BEVERIDGE, S., AND C. R. NELSON (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to the Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, 7, 151-174.

BLANCHARD, O., AND D. QUAH (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 653-658.

BRY, G., AND C. BOSCHAN (1971): "Cyclical Analysis of Time Series: Selected Procedures and Computer Programs," Technical Paper, n. 20. New York: National Bureau of Economic Research.

Burns, A. F., and W. C. Mitchell (1946): *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Evans, G., and L. Reichlin (1994): "Infomation, Forecast and Measurement of the Business Cycle," *Journal of Monetary Economics*, 33, 233-254.

Granger, C. W. J., and R. F. Engle (1987): "Co-integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251-276.

Hamilton (1989): "A New approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357-384.

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

Hodrick, R. J., and E. C. Prescott (1997): "Postwr U.S. Business Cycles: an Empirical Investigation," *Journal of Money, Credit, and Banking*, 29, 1-16.

Laxton, D., and L. Tetlow (1992): "A Simple Multivariate Filter for the Measurement of Potential Output," *Technical Report*, Bank of Canada, 59, June.

Orphanides, A., and S. van Norden (2002): "The Unreliability of Output Gap Estimates in Real Time," *The Review of Economics and Statistics*, 84, 569-583.

Persons, W. M. (1919): "Indices of Business Conditions," *The Review of Economics and Statistics*, 1, 5-110.

Stock, J. H, and M. W. Watson (1991): "A Probability Model of the Coincident Economic Indicators," in *Leading Economic Indicators: New Approaches and Forecasting Records*, K. Lahiri and G. Moore Eds., Cambrige University Press, Cambridge.

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

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Detrending Techniques"

THE IMPACT OF THE MACROECONOMIC HYPOTHESIS ON THE ESTIMATION OF THE OUTPUT GAP USING A MULTIVARIATE HODRICK-PRESCOTT FILTER: THE CASE OF THE EURO AREA

By Odile Chagny and Matthieu Lemoine

OFCE, Département Analyse et Prévision

The paper proposes to apply the multivariate Hodrick-Prescott (HPMV) filter to the estimation of the Euro area output gap. It investigates specifically the impact of using alternative economic relationships on the output gap estimates of the Euro area. The comparison with the univariate HP filter shows that this can significantly modify the appreciation of the output gap level of the Euro area in some specific periods. The paper proposes also to estimate the HPMV models with the methodology adopted for the estimation of state-space models. This strategy provides an alternative to the calibration of the parameters and allows to assess the reliability of the HPMV output gap estimates. Estimated weights associated to the economic relationships in the optimisation program of the HPMV are generally coherent with the calibrated values usually retained in the literature. The assessment of the reliability of the alternative output gaps and of the revision properties shows a substantial superiority of some HPMV models over the univariate HP filter. Finally, integrating macro-economic information improves generally the accuracy of the inflation forecasts.

1 Introduction

The concepts of potential GDP and output gap are widely used in macroeconomics even though their definition and estimation raise a number of theoretical and empirical questions, which reflect the ongoing controversy over the causes of economic fluctuations. Indeed, since the seminal contribution of Nelson and Plosser (1982) suggesting that output series are better described as integrated stochastic processes, measuring the permanent component of output, e.g. potential output, with any degree of accuracy has proved to be difficult.

An output gap is defined as the difference between - unobservable - potential and actual GDP. The precise understanding of the output gap concept therefore requires a precise definition of potential output. Potential output is commonly defined as the "maximum output an economy can sustain without generating a rise in inflation" (De Masi, 1997) or equivalently as the level of the output consistent with a stable inflation rate given the productivity shock of capital. Hence, a level of real output above potential output, i.e. a sustained positive output gap, indicates demand pressures and signals to the monetary authority that inflation pressures are increasing and that policy tightening may be required. This implies, in turn, that an accurate measure of potential output presents an important challenge to policy makers. In this respect, since the contribution of Okun (1962), numerous methodologies have been proposed in the literature.

One approach consists of using non-structural univariate methods, which separate a time series into a permanent and a cyclical component. In this context, one may distinguish between deterministic detrending, mechanical filters and unobserved components models. A first approach includes standard linear, phase average or robust detrending estimations of potential output. Such methods are obviously overly simplistic and do not capture the overall data generating process. A second methodology uses mechanical filters such as the Hodrick-Prescott (HP) filter or the band-pass filter developed by Baxter and King (1995). Nevertheless, the limits of such filters are now well-known. A third methodology assumes that output contains an unobserved component and a temporary component consisting of a random walk with drift and a stationary autoregressive process. However, Quah (1992) has shown that "without additional *ad hoc* restrictions those (univariate) characterisations are completely uninformative for the relative importance of the underlying permanent and transitory components". Another strategy for identifying the potential output involves non-statistical structural methods. In this case, the detrending method relies on a specific economic theory as for instance, the production function-based approach. This has the advantage of identifying explicitly the sources of growth, but raises a number of issues, such as the choice of an appropriate production function or the measurement of unobservable variables (total factor productivity).

Partly in response to criticisms of the previous methodologies, a variety of multivariate methodologies have been proposed as an alternative. The main rationale behind these methodologies is that adding economic content can improve the trend-cycle decomposition, through a better disentanglement of supply and demand shocks. In this paper, we restrict our attention to one of these methods, the multivariate Hodrick-Prescott filter (HPMV). This filter is a middle ground between univariate statistical methods and full, simultaneous estimations of potential output and consists merely of using additional economic relationships and/or terminal constraints in the HP maximisation program. This paper proposes to apply the HPMV to the estimation of the euro area output gap. It investigates the impact of using alternative economic relationships on the measure of the output gap of the euro area and compares the estimates with the output gap obtained with the univariate HP filter. It proposes to estimate the various HPMV models with the methodology adopted for the estimation of state-space models. This strategy allows to propose an alternative to the calibration of the parameters and to obtain confidence intervals for the output gap.

The paper is organised as follows. The second Section explains the main features of the HPMV filter. The third Section presents the estimation methodology. The models chosen for the estimation of the Euro area output gap are presented in a first part. The estimation procedure is explained in a second part. The fourth Section presents the empirical results of the application of the HPMV models to the estimation of the euro area output gap. The first part presents the data set. The second part analyses the main properties of the output gap estimates. The third part assesses the results on the basis on different criteria (standard error of the output gaps, comparison of quasi real time estimates with two sided estimates, predictive power of the output gap estimates relative to inflation). The fourth part analyses the impact of using alternative weights for the economic relationships. The fifth part concludes.

2 General description of the HPMV filter

2.1 The minimisation program

The multivariate Hodrick-Prescott filter was firstly developed by Laxton and Tetlow (Laxton and Tetlow, 1992). It is formally an extension of the optimisation problem that defines the HP filter. In the univariate version of the HP filter the trend output is the solution of a program that minimises the variance of the cyclical component subject to a smoothness constraint penalising the variations in second difference of the growth rate of the trend:

$$y^*_{t=1:T} = \arg\min(\sum_{t=1}^{T}(y_t - y^*_t)^2 + \lambda\sum_{t=2}^{T-1}\left[\left(y^*_{t+1} - y^*_t\right) - \left(y^*_t - y^*_{t-1}\right)\right]^2), \lambda > 0 \qquad (1)$$

where y_t is the output, y^*_t is the trend output, λ is the smoothness parameter. The greater is λ, the greater is the proportion of the changes in the series attributed by the filter to demand shocks; the smaller is λ, the greater is the proportion of the series attributed by the filter to supply shocks. The standard value of λ is 1600 for quarterly data. King and Rebello (1993) show that the optimal value of λ is a function of the ratio of demand and supply shocks $\sigma^2_{(y_t - y^*_t)}/\sigma^2_{\Delta^2 y^*_t}$. The problem is however that there is considerable uncertainty about the relative proportion of demand and supply shocks and hence about the optimal value of λ. Moreover, as the HP filter with a standard value of 1600 attributes "any dramatic change in the series to both supply and demand influence" (Laxton and Tetlow, 1992), it will have difficulties to distinguish between persistent demand shocks and supply shocks, as the only identifiable difference between them for a univariate filter is that the former have permanent effects, while the latter have only temporary effects.

The HP filter is a two sided filter in the middle of the sample, but becomes one-sided at the end of the sample. Consequently, the problem of accurately distinguishing between supply and demand shocks is increased near the end of the sample period. As future data is not available, the filter is less reluctant to consider that a transitory shock will revert and hence will produce unstable trend estimations.

The HPMV addresses these issues in three ways. The first is to incorporate additional economic information about excess supply or excess demand. The second is to give more flexibility to the decomposition between supply and demand shocks via allowing the weights of the smoothing parameters to take different values over time. The third is to introduce additional constraints, aimed at improving the end of sample properties of the univariate HP filter.

A general representation of the HPMV program is:

$$\begin{aligned} y^*_{t=1:T} = \arg\min \ & \left(\textstyle\sum_{t=1}^{T} (y_t - y^*_t)^{\,2} \right. \\ & +\lambda \textstyle\sum_{t=2}^{T-1} \left[\left(y^*_{t+1} - y^*_t\right) - \left(y^*_t - y^*_{t-1}\right)\right]^{\,2} \\ & \left. + \textstyle\sum_{i=1}^{n} \sum_{t=1}^{T} \mu^i_t (\varepsilon^i_t)^2 + \sum_{j=1}^{m} \sum_{t=1}^{T} \gamma^j_t (AD^j_t)^2 \right) \end{aligned} \tag{2}$$

where y_t is output , y^*_t is trend output, λ, μ^i_t and γ^j_t are the positive weights associated to the different objectives, ε^i_t are the residuals of the $i = 1$ to n economic relationships linking the output gap or the trend itself to other economic variables and AD^j_t are the additional constraints. The trend is hence chosen to simultaneously minimise the deviations of output from trend, the changes in the trend growth rate, fitting best the n economic structural relationships, conditional to j additional constraints.

As in the univariate version of the filter, the trend-cycle decomposition provided by the HPMV assumes that the output is a white noise, that the growth rate of the trend is a random walk (the output is hence I(2)), and that innovations to the cycle and the trend are orthogonal. The economic relationships amend the univariate filter as they integrate additional information to the trend-output decomposition.

2.2 Description of the additional objectives

The choice of the economic relationships is above all motivated by their information content with regard to excess supply or excess demand in the economy. The relationships most frequently found in the literature are the inflation-output relation (3), the Okun's law (4) and the capacity utilisation equation (5).

The inflation equation accounts for the view that potential output is defined as the maximum level of output that can be produced without creating inflation pressures.

The Okun's law accounts for the fact that labour market conditions contains valuable information about desiquilibrium in the goods and service market.

Finally, the use of the capacity utilisation rate provided by surveys accounts for the idea that the potential output is in the short run limited by the capital stock available in the economy.

$$\Pi_t = \Pi^e_t + a\,(L)\,(y_t - y^*_t) + \sum_{i=1}^{n} \Phi^\pi_i\,(L)\,X^\pi_{i,t} + \varepsilon_{\pi,t}, \tag{3}$$

where Π_t is the inflation, Π_t^e the anticipated inflation, $X_{i,t-1}^{\pi}$ are additional variables explaining the inflation.

$$UNR_t - UNR_t^* = b(L)\left(UNR_{t-1} - UNR_{t-1}^*\right) + c(L)\left(y_t - y_t^*\right) + \sum_{j=1}^{m} \Phi_j^u (L) X_{j,t}^u + \varepsilon_{u,t}, \quad (4)$$

where UNR_t is the unemployment rate, UNR_t^* is the equilibrium unemployment rate, $X_{j,t}^u$ are additional variables explaining the unemployment gap.

$$CAP_t - CAP_t^* = d(L)\left(CAP_{t-1} - CAP_{t-1}^*\right) + e(L)\left(y_t - y_t^*\right) + \varepsilon_{cap,t}, \quad (5)$$

where CAP_t is the capacity utilisation rate provided by surveys and CAP_t^* the equilibrium value of the capacity utilisation rate.

Equilibrium values of unemployment and capacity utilisation are exogenous. The amount of structural relationships integrated into the filter and their specification depends on the judgment or the objectives of the researchers. In the original version of the HPMV (Laxton and Tetlow 1992), two economic relations were integrated (the inflation equation and the Okun's law). Conway and Hunt (1997) and de Brouwer (1998) add the capacity utilisation rate equation. Haltmaier (1996) incorporates the inflation equation only, whereas Côté and Hostland (1993) integrate three different measures of inflation, along with an unemployment equation. In most of the specifications of the inflation equation, anticipated inflation is approximated combining lagged values of inflation and the relation is linear and symmetric[1].

The additional variables in the inflation equation account for the idea that temporary shocks do not affect the potential output. Such variables can be found in Haltmaier (1996) and in Côté and Hostland (1993). With regard to the Okun's law, the alternative specifications account for different views about the relative adjustments of labour and goods markets. Deviations between output and its potential value is for instance assumed to follow deviations on the labour market after four quarters in Conway and Hunt (1997), whereas Laxton and Tetlow (1992) assume that labour market disequilibrium adjust slowly to the goods market disequilibrium.

The idea of allowing weights to take different values over time to accounts for the researcher's judgement of supply and demand shocks for specific periods was introduced by Laxton and Tetlow (1992), whereas the idea to integrate additional restrictions to the HPMV aimed at improving the end of sample properties of the filter was first proposed by Butler (1996). The first additional restriction proposed by Butler (1996) is a steady-state condition regarding the end-of-sample estimates. To this end, an additional term is integrated in the minimisation program that penalises the variations of the trend component or the difference between the growth rate of the trend and a steady state growth rate. The second additional constraint proposed by Butler (1996) is a recursive updating restriction, that penalises the deviation between the trend and its previous quarter's estimate. Whereas the first restriction intends to prevent large, quarter-over-quarter revisions to the trend at the end of the sample, the second restriction intends to prevent large revisions to the level of the trend with the arrival of new data (Butler, 1996).

[1] However, expected inflation is modelled as an average of survey/expert's forecasts and of past inflation in Conway and Hunt (1997) and in Butler (1996), and the inflation equation is asymmetric in Butler (1996), accounting for the view that the cost of reducing inflation may be higher than the benefits of an excess demand. In both cases, these choices reflect the wish to estimate potential output consistent with the macro-economic models they are used in conjunction with.

2.3 Key implementation problems

Two key problems arise for the implementation of the filter. The first is that the estimation of the structural equations requires to know the potential value of one or more macro-economic variables, whereas the optimisation programme depends itself on the estimated residuals of these economic relationships.

The second problem is the determination of the weights λ, μ_t^i ($i = 1$ to n) associated to the different objectives of the filter. Both problems are correlated, as the value of potential output depends itself on the weights.

2.3.1 Calibration versus estimation of the parameters of the structural equations

Different solutions to the circularity problem are available in the literature.

Recursive procedure A first - and widely used - possibility is to use an iterative estimation procedure. An initial estimate of the potential is computed using an univariate HP filter and integrated in the economic relations. Residuals of the structural relations are then integrated in a second step into the optimisation program of the HPMV, etc. But in many cases, structural relations are partially or totally calibrated, using either well-established empirical studies or in coherence with macro-economic models[2]. Moreover, although the optimisation program can be generalised to produce the potential values of the other economic variables intervening in the structural equations[3], many empirical studies use exogenous estimates of these variables[4]. This estimation procedure presents some shortcomings. It is not straightforward, as many iterations are necessary[5]. It does not provide confidence intervals for the trend output. Finally, the quality of the fit of the economic relationship is not guaranteed, as they are mostly calibrated.

Kalman filter procedure Recent applications of the Kalman filter techniques provide some alternatives to overcome the difficulties of the procedure described above. Harvey (1985) explains how to reproduce the univariate HP filter with the Kalman filter. Estimating the NAIRU for France and the United-States, Boone (2000) extends this work to reproduce the HPMV with a Kalman filter[6]. This estimation procedure allows a joint estimate of the parameters of the inflation equation and the NAIRU and also provides confidence intervals for trend output. For these reasons, we have decided to use such a procedure for the estimation.

[2]Conway and Hunt (1997) constraint all coefficients of both the Okun's law and the capacity utilisation equation and estimate only the coefficients of the inflation equation. In contrast, Butler (1996) calibrates all structural equations, whereas all coefficients are estimated in Laxton and Tetlow (1992) , Haltmaier (1996) and de Brouwer (1998).

[3]See Haltmaier (1996).

[4]Laxton and Tetlow (1992) estimate the NAIRU using a multivariate HP filter, whereas Conway and Hunt (1997) use an univariate HP filter for the estimation of the NAIRU and the average value over the sample period for the capacity utilisation rate (allowing for a break in the series). De Brouwer (1998) assumes that the NAIRU stems from a structural estimation, which leads persistent and substantial unemployment gaps.

[5]For instance, 250 iterations are necessary in de Brouwer (1998).

[6]For recent applications, see also Lemoine and Pelgrin (2002).

2.3.2 Choosing the weights

The determination of the weights involves the choice of the smoothing parameter λ, the choice of the weights of the economic relationships, and the stability of the weights over time. In most empirical studies, the value of the smoothing parameter λ is fixed at its conventional value (1600 in the case of quarterly data)[7], imposing an *a priori* assessment on the relative smoothness of the trend and the cycle.

The determination of the weights of the economic relationships is identified as one of the key questions in implementing the HPMV and has received no clear and definitive answer. One solution attempted is to test alternative weighting and to see how sensitive the results are to these choices[8]. An alternative[9] is to link the weights of the economic relations to their uncertainty. In many cases, the decision rules are however more or less arbitrary[10], even with Kalman filter procedures. The same conclusions can be drawn for the choice of the weights associated to the additional constraints[11]. The strategy adopted here has been to endogenize the determination of the weights of the economic relationships in the HPMV optimisation program, using their estimated variances.

3 Estimation methodology

3.1 The models chosen for the estimation of the euro area output gap

In this paper, we concentrate on the impact of integrating alternative macroeconomic hypotheses on the estimation of the output gap of the euro area. The HPMV framework.developed here incorporates hence no additional constraints and weights are assumed to be time invariant.

Seven models have been estimated. The first model is the univariate HP filter. The second model, named HP-P, incorporates an inflation equation (6):

$$\Pi_t = \alpha_1 \Pi_{t-1} + \alpha_2 \Pi_{t-2} + \alpha_3 \Pi_{t-3} + \alpha_\pi \left(y_{t-1} - y^*_{t-1} \right) + \gamma_1 S_{t-1} + \gamma_2 S_{t-2} + \varepsilon_{\pi,t} \quad (6)$$

[7] A noticeable exception is Côté and Hostland (1993). Their statistical methodology is indeed to choose the value of the smoothing parameter λ that maximizes the likelihood of their system of structural equations. To this end, values of λ ranging from 10 to 100 000 are tested. They find that the values are sensitive to the sample period and that the results are sensitive to the specification of the structural equations.

[8] This approach is used by Laxton and Tetlow (1992) (essentially through dropping one economic relation) and Haltmaier (1996).

[9] This approach is mentioned by Laxton and Tetlow (1992), tested by Butler (1996) and used by De Brouwer (1998). In the latter, the sum of the squared residuals for each of the conditioning equations is weighted by its size relative to the sum of squared residuals of the output gap from the previous iteration and weights change at each iteration.

[10] Haltmaier (1996) compares for instance the impact of different weights on the fit of the inflation equation and the standard deviation of the output gap, but admits that this criteria is not well grounded.

[11] In the Butler model, the weights put on the additional constraints are fixed at the positive value 64 for the final 15 quarters of the sample, and are at the value zero otherwise. A steady- state growth rate restriction is also found in Conway and Hunt (1997). The weights are fixed to one, except for the last quarter, where the value is zero.

where Π_t is the inflation, measured as the quarterly growth rate of the consumer price index, S_t is the relative import price growth rate, measured as the difference between the quarterly growth rate of the goods and services deflator index and the quarterly growth rate of the consumer price index. The choice of a relative price index has been motivated by the idea that supply shocks entering the inflation equation are assumed to revert to zero in the long run. In order to ensure that the inflation equation is in the long run specified in acceleration, the sum of the coefficients α_1 to α_3 has been constrained to 1. Different lags for the output gap have been tested, but in most of the estimated models, the output gap showed the highest explanatory power at lag 1 and first difference of the output gap was found not significant.

The third model, named HP-P-CAP, incorporates the inflation equation and a capacity utilisation equation specified as follows:

$$CAP_t - CAP_t^* = \alpha_{cap} \left(y_t - y_t^* \right) + \varepsilon_{cap,t}, \tag{7}$$

where CAP_t^* has been estimated as the linear trend of the capacity utilisation rate, accounting for the view of possible structural changes in the trend level of the capacity utilisation rate in the euro area.

The fourth and the fifth models, named HP-P-OK1 and HP-P-OK2, incorporate the inflation equation and an Okuns' law specified as follows:

$$UNR_t - UNR_t^* = \alpha_{unr} \left(y_{t-1} - y_{t-1}^* \right) + \varepsilon_{u,t}, \tag{8}$$

where two alternative exogenous measures of UNR^* have been used (an univariate HP trend and a time-varying Nairu (see below for a more detailed description of the TV Nairu).

The sixth and the seventh models, named HP-P-OK1-CAP and HP-P-OK2-CAP, incorporate three economic relationships : the inflation equation, the Okun's law and the capacity utilisation equation. The model HP-P-OK1-CAP uses as exogenous UNR^* the TV Nairu, whereas the model HP-P-OK2-CAP uses as exogenous UNR^* the trend unemployment provided by an univariate HP filter.

3.2 Estimation procedure

The estimation procedure for the general HPMV filter consists firstly in rewriting the minimisation program of the HPMV filter in a state-space form. The state-space form of the general HPMV filter with inflation (Π), unemployment rate (UNR) and capacity utilisation rate (CAP) can be written [12] as follows:

$$\left\{ \begin{array}{ll} SV_t = A.SV_{t-1} + U_t & \text{(state equation)} \\ MV_t = B.SV_t + C.EV_t + V_t & \text{(measurement equation)} \end{array} \right. \tag{9}$$

with $U \sim NID(0,Q)$, $V \sim NID(0,R), Cov(U,V) = 0$[13] and with:

[12] Numbers of lags for inflation, output gap and exogenous variables have been fixed with usual t-statistics and AIC criteria.

[13] A covariance equal to 0 corresponds to the classical assumption that the state-space model is formulated in its canonical form.

- the state vector: $SV_t = (Y_t^*, Y_{t-1}^*, gap_t, gap_{t-1})'$
- the measurement vector: $MV_t = (Y_t, \Pi_t, UNR_t, CAP_t)'$
- the exogenous vector: $EV_t = (\Pi_{t-1}, \Pi_{t-2}, \Pi_{t-3}, S_{t-1}, S_{t-2}, UNR_t^*, CAP_t^*)'$

and matrices

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & \alpha_\pi \\ 0 & 0 & 0 & \alpha_{unr} \\ 0 & 0 & \alpha_{cap} & 0 \end{bmatrix},$$
$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 & \gamma_1 & \gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{10}$$
$$U_t = \begin{bmatrix} \varepsilon_{\Delta\Delta y^*,t} \\ 0 \\ \varepsilon_{gap,t} \\ 0 \end{bmatrix}, V_t = \begin{bmatrix} 0 \\ \varepsilon_{\pi,t} \\ \varepsilon_{unr,t} \\ \varepsilon_{cap,t} \end{bmatrix}$$

In the state-space framework presented here, the calibration of the parameter λ and of the weights μ_i associated to the economic relationships consists in fixing the following variance ratios:

$$\lambda = \sigma^2_{gap}/\sigma^2_{\Delta\Delta y^*}, k_\pi = \sigma^2_{gap}/\sigma^2_\pi, k_{unr} = \sigma^2_{gap}/\sigma^2_{unr}, k_{cap} = \sigma^2_{gap}/\sigma^2_{cap}$$

However, if we except λ[14], which has been imposed here at its conventional value of 1600, these ratios can be estimated in the state-space framework, which allows to compare the results obtained with these estimated ratios with the results obtained with the usual values found in the literature.

In a second step, the state variables (Y_t^*, gap_t) and parameters

$$\left(\alpha_\pi, \alpha_{unr}, \alpha_{cap}, \alpha_1, \alpha_2, \alpha_3, \gamma_1, \gamma_2, \sigma^2_{\Delta\Delta y^*}, \sigma^2_\pi, \sigma^2_{unr}, \sigma^2_{cap}\right)$$

are estimated with state-space estimation techniques, i.e. the Kalman and EM algorithms described in appendices 6.1 and 6.2.[15] The EM algorithm is an iterative algorithm designed to provide maximum likelihood estimates of the parameters. Each step of the EM algorithm is itself decomposed into two steps that update estimates of state variables and parameters, conditional to each other. The update of the state variables conditional to parameters is implemented with a Kalman filter, whereas the update of parameters conditional to state variables is obtained by the maximisation of the conditional likelihood. These two steps are repeated untill convergence towards the maximum likelihood estimates is reached.

However, the implementation of the algorithms requires to face two problems. At each step of the EM algorithm, the Kalman filter has to be initialised with appropriate initial values and

[14]If the cycle is specified as a white noise, λ has to be fixed to 1600 for quarterly data. Indeed, this parameter would be estimated to 0, when such a constraint is relaxed. A correct estimation of this variance ratio would require a correct AR(2) specification of the cycle and the model would exit from the HPMV framework.

[15]For a comprehensive presentation of the estimation of state-space models, see Harvey (1989), Durbin and Koopman (2001).

variances of the state variables. As the global EM algorithm for maximising the likelihood can be trapped in local minima, *a priori* knowledge of parameters' distribution has to be used. These *a priori* values are incorporated in the initialisation of the algorithm to help its convergence. The following strategy has been adopted here:

- *Initialisiation of the Kalman filters*: initial possible values and variance of the output gap have been taken from the univariate HP filter. As the trend and the lagged trend are integrated, few *a priori* are available on their location and a diffuse initialisation is necessary[16]. Initial trend values have been estimated with the univariate HP filter and an arbitrary high value has been chosen for their variance.

- *EM initialisation*: elasticities and residual variances have been initialised by pre-estimating equations (6), (8) and (7) with the univariate HP output gap.

Together with the descriptive statistics of estimated output gaps, state-space models allow to compute various interesting statistics. Confidence bands around the output gap can be measured taking into account "filter" and "parameter" uncertainty, following Hamilton (1986) procedure (Appendix 6.3).

Given the well-known end-of-sample problems of the univariate HP filter, output gap revisions should also be considered. Following Orphanides and Van Norden (1999) typology, state-space models allow to compute "quasi-real time", "quasi-final" and "final" output gap estimates. The quasi-real time estimates use subsamples of final data[17]. Quasi-final (filtered) estimates are based on sub-samples estimates of the state variables, conditional to the full sample parameters estimates. Lastly, the final (smoothed) estimates use the full sample for the estimation of parameters and state variables. The difference between quasi-final and quasi-real series reflects the importance of the parameter instability in the underlying state-space model, while the difference between quasi-final and final series reflects the importance of *ex post* information in estimating the output gap.

Finally, as HPMV models combine output gap and inflation dynamics, estimated state-space models provide forecasts of inflation, conditional to output gap estimates. The comparison of the predictive accuracy of HPMV output gap estimates with benchmark forecasts can hence be tested with Diebold-Mariano tests (Appendix 6.4 for a description of the test).

4 Application to the Euro area

4.1 Data set

Estimations are made for the Euro area 12. The estimation period covers the years 1980 to 2000 and is made on a quarterly basis. The choice of this estimation period has been mainly motivated by the wish to use ESA95 GDP data for most of the countries.

[16] Diffuse initialisation of non-sationary variables is presented and explained in Durbin and Koopman (2001).
[17] They do not account for data revisions, as would be the case with real time estimates.

The following data set has been used. Inflation data use the harmonised price index published by Eurostat. For the years 1980-1990, they have been backward calculated on the basis of national price indexes. Seasonal adjustment has been implemented with X12 Arima. The capacity utilisation rate published by the European Commission has been used from 1992 on. Earlier data have been backward calculated using data published by the European Commision for the European union (EU15). The quarterly GDP data (constant prices of 1995) come from Eurostat National accounts. They have been backward calculated with the National accounts when necessary. German post reunification data have been linked to the west German National accounts data, in order to avoid the break of 1991 due to the reunification. Unemployment rate is the standardized unemployment rate published by Eurostat from 1997 on. Earlier values come from the data set calculated by Fagan et al. (2001), covering the same twelve countries. The exogenous Nairu and the import price deflator come from the same data set.

4.2 Main properties of output gap estimates

Four main results emerge from the comparison of the descriptive statistics relative to the estimated models and output gaps.

Firstly, it appears difficult to obtain a significant impact of the output gap on the inflation in the bivariate model HP-P, which integrates as additional economic relationship the inflation equation only (see Table 1a). This is mainly due to the fact that the estimation procedure chosen here, which consists in estimating all parameters in an integrated framework -instead of calibrating them-, provides in this specific case output gap estimates which are hardly different from the univariate HP filter (see Graphs 1 to 6). This result confirms the relative poor performance of the univariate HP filter to contain information with regard to inflation[18]. Table 1a highlights hence the advantage of multivariate models, as in all cases the integration of other additional economic relationships provides better fits of the inflation equation. This is especially true when - in addition to the inflation equation - the HPMV integrates the capacity utilisation rate (model HP-P-CAP), or when it links the output gap to an unemployment gap which incorporates itself some information about inflation pressures, as is the case with the models HP-P-OK1 and HP-P-OK1, using as exogenous measure of the equilibrium unemployment rate a TV Nairu estimated within a Gordon (1997) framework[19].

Secondly, the results highlight the relative advantage of using a coherent framework for estimating all coefficients and variables entering the HPMV filter. This is illustrated by the comparison of the four models integrating the Okun's law equation. Models HP-P-OK1 and HP-P-OK1-CAP use as exogenous variable Nairu estimates obtained with a multivariate unobserved component model integrating an inflation equation. The Nairu is very smooth, which leads to very persistent and high unemployment gaps[20], as in other empirical studies estimating HPMV output gaps (see for instance de Brouwer (1998) in the case of Australia). As a consequence, the estimated output gaps exhibit persistent periods of high excess supply, if one

[18]See Chagny and Döpke (2001) for an empirical evidence for the euro area, covering the same period of time.

[19]Unpublished estimates show that estimating the Nairu with an HPMV integrating an inflation equation would also provide better results than a univariate HP filter based Nairu.

[20]Over the period 1981-2000, the effective unemployment rate is on average 1.4 percent above the TV Nairu, whereas the unemployment gap calculated with a univariate filter is on average equal to zero.

excepts the years 1990-1991 and 2000[21] . But this result has a counterpart : as a consequence of the large unemployment gap, the estimated coefficient of the output gap in the Okun's law fluctuates around -0.7, a value quite higher than one would expect for an Okun's coefficient. In contrast, the HP-P-OK2 and HP-P-OK2-CAP models use similar methodologies for estimating the Nairu and the output gap, as the Nairu is obtained here with an univariate HP filter. This leads in average to much lower unemployment gaps, but also to more coherent estimated values for the Okun's law coefficient (around -0.3, Table 1a) and to a better fit of this economic relationship. Integrating the "researcher's view" on the equilibrium value of some exogenous variables might hence have consequences on the reliability of the HPMV estimates.

Thirdly, and letting aside the models HP-P, HP-P-OK1 and HP-P-OK1-CAP described here above, the analysis of the descriptive statistics concerning the output gaps confirms the results obtained in other empirical studies. The integration of macroeconomic information within the HPMV filter can in some specific periods provide quite an "alternative" information about the level of the output gap, although the differences with the univariate HP filter are on average limited[22] . All models provide output gaps with higher standard deviation than the HP univariate filter and, except for the capacity utilisation model, less volatile trends (see Table 1b). The highest output gap standard deviation is obtained with the capacity utilisation model (HP-P-CAP), as it partly captures the higher volatility of the industrial activity. All models show more excess demand at the end of the eighties and more excess supply in the first half of the nineties. At the end of the eighties, the difference is the most pronounced for the capacity utilisation rate model, as a consequence of the sustained pressures on the industrial capacities consecutive among others to the boom caused by the iron curtain fall on the German economy. The appreciation of the actual output gap level, understood as the end of sample output gap estimates, differs significantly from one model to another. Models integrating Okun's law suggest that a greater part of the recovery of the end of the nineties was demand induced, which is reflected by a level of the unemployment rate below it's equilibrium value and a consequently lower growth rate of potential output (Table 1b). In contrast, the capacity utilisation based model exhibits less demand excess in 2000 and the highest end of sample potential output growth rate, reflecting the impact of the investment recovery on the capacity utilisation rate.

Fourthly, the analysis of the estimated weights provides some explanations for the comparison of the estimated output gaps and benchmarks for a comparison with other empirical studies. In most cases, it appears that the estimated weights of the economic relationships are coherent with the calibrated values retained in most empirical studies (Tables 1c and 1d). The consequences of using alternative weights on the output gap estimates will be assessed in the fourth part of this Section. The highest weight for the inflation equation is obtained with the TV Nairu models, who exhibit both a high volatility of the output gap and a good fit of the inflation equation. Similar conclusions, although to a lesser extent, can be drawn for the capacity utilisation model (HP-P-CAP). Finally, the very similar output gap estimates obtained with the models HP-P-OK2 and HP-P-OK2-CAP are explained by the fact that the estimation procedure gives a very low weight to the capacity utilisation equation in the latter model.

[21]Moreover, the unit root test for the output gap is not rejected for these both models, in contrary to all others models (see table 1b).

[22]The difference between the absolute value of the models HP-P-CAP, HP-P-OK1, HP-P-OK1-CAP output gaps and the absolute value of the univariate HP filter output gap amounts on average 33 percent of the standard deviation of the univariate HP filter output gap.

4.3 Assessment of output gaps estimates

4.3.1 Uncertainty statistics

Table 2.a directly illustrates the well-known end-of-sample problem of the univariate HP filter. This filter cannot inform the policy-maker, as the standard errors indicate that the filtered gap is never significantly different from zero and that standard revisions are very high.[23] Although results become acceptable for the smoothed gap, there is still important uncertainty about the "true" value of the output gap.[24] Univariate HP filters have difficulties to distinguish between persistent demand shocks and supply shocks, as the only identifiable difference between them for a univariate filter is that the former have permanent effects, while the latter have only temporary effects. We might hence expect to improve one-sided results with multivariate filters, i.e. to better distinguish supply from demand shocks with additional economic information. It is also of some interest to see whether the addition of exogenous variables can also statistically improve the accuracy of the output gap estimates. Both issues are explored here[25].

The results presented in Table 2.a show that the integration of exogenous variables decreases the filter uncertainty of the two-sided estimates. The lowest filter uncertainty is obtained for the multivariate models with Okun's law. However, models using the TV-Nairu (HP-P-OK1 and HP-P-OK1-CAP) show a higher parameter uncertainty, which leads to increase the global standard errors relative to the univariate HP filter.

Concerning the one-sided estimates, the comparison of the standard errors shows that all multivariate models perform better than the univariate HP filter. This is especially the case for the capacity utilisation model (HP-P-CAP). For multivariate models (except HP-P), all confidence bands widths are lower than the standard deviations. Multivariate filtered gaps are hence generally significantly different from zero.

Contrary to the univariate HP filter, the standard deviations of revisions of all multivariate HP models (one sided versus two-sided estimates) are lower than the standard deviations of the output gaps. The lowest revisions occur with the Okun's law based models. In conclusion, the addition of capacity utilisation or Okun's law allows to better distinguish between supply and demand shocks, even with filtered estimates. However, these results are downward biased, as they do not take account of possible revisions of the Nairu estimates. The impact of such revisions on quasi-real time revisions is analysed in the next part.

4.3.2 Revision statistics

In order to better account for the end of sample properties of the models, "quasi-real time" output gaps have been estimated for four models[26]. This allows to assess the global impact of

[23] The filtered gap, with a standard deviation of 0.94, has a confidence band of 2*54=1.08 and a standard revision equal to 1.09.

[24] The confidence band width, with a value of 0.56, is lower than the standard deviation equal to 0.83.

[25] Following Hamilton (1986) procedure, these standard errors are decomposed into "parameter" and "filter" uncertainty (Appendix 7.3).

[26] As quasi-real time estimates of TV-Nairu were not available, models HP-P-OK1 and HP-P-OK1-CAP have been excuded from this revision study.

parameters instability and of revisions to the univariate HP Nairu for the models HP-P-OK2 and HP-P-OK2-CAP. Recursive estimations have been implemented from 1997:1 to 2000:3. Quasi-real time estimates have not been computed before 1997, in order to keep a sufficiently large sample for the state-space estimation.

Tables 2.b and 2.c provide essentially two types of results (mean and standard deviation of revisions), which lead to similar interpretations.

Firstly, the output gaps exhibit positive means in quasi-real time and means near to zero in final time estimates. This is directly related to the recovery of the European economy in the years 1998-2000. However, this effect is low for the HP-P-CAP model and is particularly strong for the Okun's law based models (HP-P-OK2 and HP-P-OK2-CAP).

Secondly, the standard deviations of the revisions are in general higher than for the univariate HP filter in the case of the Okun's law based models (indicators NS in Table 2.c). However, the model HP-P-CAP performs quite well, better than the univariate HP filter : the HP-P-CAP gap would be estimated in quasi-real time with a better standard deviation than all other models. The bad relative performance of the Okun's law based models can be related to the impact of the Nairu revisions which were not taken into account in the previous part.

4.3.3 Inflation forecasts

All estimated multivariate HP models relate inflation to the lagged output gap. Thus, their estimated state-space form can provide forecasts of inflation, conditional to the output gap smoothed estimates.

To assess such inflation forecasts, the best way consists in comparing their Root Mean Squared Errors (RMSE) with the RMSE of naive forecasts. Diebold-Mariano test (see Appendix 6.4) allows to compare the predictive accuracy of the HPMV filter and of naive models of inflation. RMSE, RRMSE, Diebold-Mariano statistics and their associated probability are presented for all HPMV models in comparison with two naive benchmarks models in Table 2.d.

When the naive benchmark model is a simple random walk model, all HPMV filters estimates provide inflation forecasts which are significantly improved : except for the HP-P model, the null hypothesis of equal predictive power is rejected at a 5% threshold.

Auto-regressive models are more difficult to beat. Consequently, the predictive accuracy of the models is in most cases not significantly different from the predictive accuracy of the auto-regressive model. The quite good predictive accuracy of the models HP-P-OK1 and HP-P-OK1-CAP may be due to the fact that information about inflation is already integrated in the TV-Nairu estimates.

In conclusion, the HMPV estimated output gaps do not exhibit significant inflation predictive accuracy. This result is probably related to the low weights of the inflation equations. The impact of alternative weights on the quality of inflation forecast is analysed in the next part.

4.4 Impact of varying weights

Tables 3.a , 3.b and the Graph 7 provide the main results associated to the use of alternative weights of the inflation equation in the case of the HP-P model. Increasing the weight of the inflation equation in the HPMV optimisation program leads to account for more excess supply during the periods of decelerating inflation (like in the years 1982-1986 and 1992-1996). Symmetrically, periods of accelerating inflation (like in the years 1988-1991 and 1999-2000) are associated to more excess demand. This result is caused by the fact that imposing high values for the weight of the inflation equation leads to a better fit of the inflation equation, associated to a higher value of the coefficient of the output gap in the equation. This information is corroborated by better inflation forecasts of the models based on high weights values of the inflation equation. For instance, the marginal probability associated to the Diebold-Mariano statistics that the HPMV has the same predictive accuracy than the auto regressive model takes the value 6% with a weight of 25, i.e. the lowest value of all models presented in this paper. But the gain with regard to inflation forecasts is obtained at the cost of higher standard errors of the output gaps, as both filter and parameter uncertainty increase when high values for the weights of the inflation equation are imposed : the standard error of the output gap amounts 0.44% of the standard deviation of the output gap with a weight of 25, to compare to 0.33% with the estimated weight value.

5 Conclusion

This paper presents for the first time estimates of the euro area output gap based on the use of alternative macroeconomic relationships within a HPMV framework implemented with state-space techniques. The comparison of the alternative output gaps estimates with the results obtained with a univariate HP filter shows that integrating macroeconomic information can in some specific periods significantly modify the appreciation of the output gap level of the euro area. For instance, all models estimated in the paper exhibit more excess demand at the end of the eighties and more excess supply in the first half of the nineties. In contrast to many empirical studies, all parameters of the models have been estimated. As a consequence, it appears that the estimated weights associated to the economic relationships in the optimisation program of the HPMV are generally coherent with the calibrated values usually retained in the literature. The assessment of the reliability of the alternative output gaps and of the revision properties shows a substantial superiority of the model based on the capacity utilisation rate over the other models. Okun's law based HPMV models lead indeed to higher quasi real time revisions than the univariate HP filter and further work is needed to endogenize the estimation of the Nairu within the HPMV framework, in order to provide a better assessment of the uncertainty of the Okun's law based output gaps and to improve the revision properties of the models. Although all models improve the accuracy of the inflation forecasts when the benchmark naive model is a random walk, the inflation forecasts results confirm the superiority of the capacity utilisation based model.

References

Baxter, M., and R. G. King (1995): "Measuring busines cycles: approximate band-pass filters for economic time series," NBER Working Paper No 5022.

Boone, L. (2000): "Alternative methods to estimate unobserved variables: the HPMV and Kalman filters approaches," OECD Economic Department Working Paper No. 240, Paris.

Butler, L. (1996): "A semi-structural method to estimate potential output: combining economic theory with a time-series filter," Bank of Canada Technical Report No. 76.

Chagny, O., M. Lemoine, and F. Pelgrin (2002): "A survey of multivariate methods for estimating the output gap," Eurostat Report, mimeo.

Chagny, O., and J. Döepke (2001): "Measures of the output gap in the Euro-zone: an empirical assessment of selected methods," Kiel Institute of World Economics Working Paper No. 1053 .

Conway, P., and B. Hunt (1997): "Estimating potential output: a semi-structural approach," Reserve Bank of New Zealand Discussion Paper 1997/09.

Côté, D., and D. Hostland (1993): "Measuring potential outupt and the NAIRU as unobserved variables in a systems framework," Economic Behaviour and Policy Choice Under Prie Stability," Proceedings of a Conference Held at the Bank of Canada, October.

De Brouwer, G. (1998): "Estimating output gaps," Reserve Bank of Australia Research Discussion Paper No. 9809.

De Masi, P. (1997): "IMF estimates of potential output : theory and practice," Staff Studies for the World Economic Outlook.

Dempster, A. P., N. M. Laird and D. B. Rubin (1977): "Maximum likelihood for incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, Series B, vol. 39, pp. 1-38.

Diebold, F. X., and R. S. Mariano (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, pp. 253-263.

Durbin, J., and S. J. Koopman (2001): *Time Series Analysis by State Space Methods.* Oxford: Oxford University Press.

Fagan, G., J., Henry, and R. Mestre (2001): "An area-wide model (AWM) for the Euro area," European Central Bank Working Paper No 42.

Gordon, R. J. (1997): "The time varying NAIRU and its implications for economic policy," *Journal of Economic Perspectives*, Vol. 11; N1, winter.

Haltmaier, J. (1996): "Inflation-adjusted potential output," Board of Governors of the Federal Reserve System International Finance Discussion Papers, No 561.

Hamilton J. D. (1986): "A standard error for the estimated state vector of a state-space model," *Journal of Econometrics*, 52, pp. 129-57.

Harvey A. C. (1985) : "Trends and cycles in macroeconomic time series," *Journal of Business and Economic Statistics*, 3, pp. 216-227.

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

Kalman, R. E. (1960): "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, Vol. 82, Series D, pp. 35-45.

King, R. G., and S. T. Rebelo (1993): "Low frequency filtering and real business cycles," *Journal of Economic Dynamics and Control*, vol. 17 No 1/2, pp. 207-231.

Laxton, D., and R. Tetlow (1992): "A simple multivariate filter for the measurement of potential output," Technical Report No. 59. Bank of Canada.

Lemoine M., and F. Pelgrin (2002): "Introduction aux modles espace-tat, filtre de Kalman et applications," Eurostat Working Paper, Forthcoming.

Nelson, C. R., and C. I. Plosser (1982): "Trends and random walks in macroeconomic time series," *Journal of Monetary Economics*, 10, pp. 139-162.

Okun, A. (1962): "Potential GDP: its measurement and significance," American Statistical Association, Proceedings of the Business and Economic Statistics Section: pp. 98-103. Washington.

Orphanides, A., and S. Van Norden (1999): "The reliabilty of output gap estimates in real-time," Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series, pp. 99-138.

Quah, D. (1992): "The relative importance of permament and transitory components : identification and some theoretical bounds," *Econometrica*, 60, pp. 107-118.

Rünstler, G. (2002): "The information content of real-time output gap estimates. An application to the euro area," ECB Working Paper.

St. Amant, P., and S. Van Norden (1997): "Measurement of the output-gap : a discussion of recent research at the Bank of Canada," Bank of Canada Technical Report No 79.

6 Appendices

6.1 Kalman filter

A classical Kalman filter (Kalman, 1960) is used to evaluate the state vector SV_t given parameters, current and past observed values of MV_t., in a state-space model of the form:

$$\begin{cases} SV_t = A.SV_{t-1} + U_t & \text{(state equation)} \\ MV_t = B.SV_t + C.EV_t + V_t & \text{(measurement equation)} \end{cases} \quad (11)$$

The current estimate and its variance are noted $SV_{t,t}$ and $\Sigma_{t,t}$. The forecast and the variance of SV at $t+1$ given MV_t are noted $SV_{t,t+1}$and $\Sigma_{t,t+1}$. Assuming such notations, each iteration of the Kalman filter can be summarised by equations 1 to 5:

1. $\widehat{SV}_{t,t} = \widehat{SV}_{t-1,t} + K_t(MV_t - D.EMV_t - C\widehat{SV}_{t-1,t})$
2. $\Sigma_{t,t} = (I - K_t C)\Sigma_{t-1,t}$
3. $\widehat{SV}_{t,t+1} = A\widehat{SV}_{t,t} + B.ESV_t$
4. $\Sigma_{t,t+1} = A\Sigma_{t,t}A' + Q_t$
5. $K_t = \Sigma_{t-1,t}C'(C\Sigma_{t-1,t}C' + R_t)^{-1}$
6. $\widehat{SV}_{-1,0} = m$ and $\Sigma_{-1,0} = P$.

The sixth equation describes initial conditions of the algorithm.

6.2 EM algorithm

The normal likelihood function can be deduced from the Kalman filter evaluation of the state vector. Parameter estimates for $\theta = (A, B, C, D, Q, R)$ are then computed by maximising the likelihood function L with respect to the unknown parameters. This maximisation is iterated with a Kalman filter through the EM strategy (Dempster et al., 1977). The EM algorithm is then an iterative algorithm, which creates a sequence of estimates $(\theta_i)_{i=1,2,\ldots}$ from an initial condition θ_0. Iterating these steps, estimates should converge towards the maximum likelihood estimators. Each iteration is decomposed into the two following steps:

- Step E : $Q(\theta, \theta_i) = E(L(y|\theta))$
- Step M : $\theta_{i+1} = \arg\max_\theta Q(\theta, \theta_i)$

6.3 Measures of uncertainty

The uncertainty around the output gap plays a major role in designing the role output gaps should play in monetary rules. Following Hamilton (1986), this uncertainty has two main sources, the parameter uncertainty and the filter uncertainty:

$$E\left(C_{t,t}(\hat{\theta}) - C_t\right)^2 = E\left(C_{t,t}(\hat{\theta}) - C_t(\theta)\right)^2 + \Sigma_{t,t}(\theta)$$

with $C^Y_{t,t}(\hat{\theta})$ the output gap estimated at t, given both the observed output until the end of the period T and the parameter estimate $\hat{\theta}$. $\Sigma_{t,t}(\theta)$ is the filter uncertainty and the left-hand term is the parameter uncertainty, noted further $\Sigma^P_{t,t}$. The filter uncertainty is directly provided at each date t by the Kalman filter. But the parameter uncertainty has to be computed separately with a Monte-Carlo simulation procedure:

- We draw K parameter sets $(\theta_k)_{k=1...K}$ with the law $N(\hat{\theta}, \hat{\sigma}^2_\theta)$
- We compute with the Kalman filter the corresponding K cycles $C_t(\theta_k)$
- We deduce for each date t a the parameter uncertainty $\Sigma^P_{t,t}$ as a mean square error criterion

$$\Sigma^P_{t,t} = \frac{1}{K} \sum\nolimits_{k=1...K} \left[C_t(\theta_k) - C_t(\hat{\theta})\right]^2$$

Moreover, we might be interested in the end-of-sample quality of the estimates, by comparing the one-sided / filtered estimates $C_{t,t}(\hat{\theta})$, that policy makers would have had at that time, and the two-sided / smoothed estimates $C^Y_{T,t}(\hat{\theta})$, that can be estimated at the end T of the period.[27] This comparison can be evaluated computing the parameter and filter uncertainty of smoothed estimates ($\Sigma_{T,t}$ and $\Sigma^P_{T,t}$). Moving to the smoothed estimates can strongly decrease the overall uncertainty.

6.4 Diebold-Mariano test

The forecasting properties of a model (M) can be assessed in comparison with a naive model (M_0), by considering the difference of their squared residuals $\hat{d}_i = \hat{\eta}^2_{M,i} - \hat{\eta}^2_{0,i}$. When the forecasted variable is assumed to be known until a final date T, these residuals $\hat{\eta}^2_{M,i}$ and $\hat{\eta}^2_{0,i}$ are estimated on a forecasting sample $[T+n+1, T+n+N]$. In this case, n is called the forecast horizon and N is the number of errors that can be used to compare the forecast accuracy of the models. For testing the null hypothesis of equal forecasting performance of two models, Diebold and Mariano (1995) proposed the statistic:

$$\begin{aligned} S_{DM} &= \frac{\bar{d}}{\sqrt{V(\bar{d})}} \xrightarrow{d} N(0,1) \\ \text{with } \bar{d} &= \frac{1}{N}\sum_{i=1}^{N} \hat{d}_i \text{ and } V(\bar{d}) = \frac{1}{N}\left(\hat{\gamma}_0 + 2\sum_{i=1}^{N}\hat{\gamma}_i\right) \end{aligned}$$

[27] Another end-of-sample uncertainty consists in future data revisions, although they are not considered here.

where $\hat{\gamma}_i, i = 0, ..., n-1$ are the estimated autocovariances of the series of the series of forecasting error differences.

Tables

Models description

Model HP: HP univariate filter.

Model HP-P: HPMV with inflation equation.

Model HP-P-CAP: HPMV with inflation equation and capacity utilisation equation.

Model HP-P-OK1: HPMV with inflation equation and Okuns'law (TV Nairu).

Model HP-P-OK1-CAP: HPMV with inflation equation, Okuns'law (TV Nairu) and capacity utilisation equation.

Model HP-P-OK2: HPMV with inflation equation and Okuns'law (HP Nairu).

Model HP-P-OK2-CAP: HPMV with inflation equation, Okuns'law (HP nairu) and capacity utilisation equation.

Main properties of the alternative output gaps

Table 1.a: Parameter estimates of the economic relationships

Models	HP	HP-P	HP-P-CAP	HP-P-OK1	HP-P-OK1-CAP	HP-P-OK2	HP-P-OK2-CAP
Inflation equation							
π_{t-1}		0.33	0.30	0.31	0.30	0.34	0.34
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
π_{t-1}		0.25	0.25	0.26	0.26	0.25	0.25
		(0.07)	(0.06)	(0.07)	(0.07)	(0.09)	(0.08)
$Y_{t-1}-Y^*_{t-1}$		0.03	0.05	0.03	0.03	0.03	0.04
		(0.28)	(0.01)	(0.02)	(0.02)	(0.14)	(0.09)
S_{t-1}		0.08	0.09	0.08	0.08	0.08	0.08
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
S_{t-2}		-0.05	-0.05	-0.05	-0.05	-0.05	-0.05
		(0.03)	(0.04)	(0.02)	(0.02)	(0.03)	(0.04)
σ^2_π		3.54e-06	3.54e-06	3.54e-06	3.54e-06	3.54e-06	3.54e-06
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
Okun's law							
$Y_{t-1}-Y^*_{t-1}$				-0.69	-0.74	-0.34	-0.34
				(0.00)	(0.00)	(0.00)	(0.00)
σ^2_{unr}				7.83e-06	9.89e-06	3.54e-06	3.72e-06
				(0.00)	(0.00)		
Capacity rate equation							
$Y_t-Y^*_t$			2.36		1.35		1.38
			(0.00)		(0.00)		(0.00)
σ^2_{cap}			1.09e-04		5.2e-06		6.13e-04
			(0.00)				
HP parameter							
$\sigma^2_{\Delta\Delta y^*}$	7.8 e-08	7.44 e-08	1.68e-07	3.84e-07	3.29e-07	8.48e-08	8.69e-08
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		

N.B.: Estimation period : 1980:1 2000.4. Into brackets: P values.

Table 1.b: Descriptive statistics for the smoothed output gap of the alternative models (1981:1 2000:4)[a]

Models	(1) Mean	(2) SD	(3) Min	(4) Max	(5) Actual growth rate of pot. output[b]	(6) SD of growth rate of pot. output	(7) Unit root test for output gap[c]
HP	0.02	0.83	-1.98	2.25	2.86	0.14	-3.49***
HP-P	-0.02	0.84	-1.98	2.23	2.82	0.14	-3.36***
HP-P-CAP	-0.09	1.14	-2.50	2.15	2.88	0.22	-2.44**
HP-P-OK1	-1.89	1.46	-3.90	1.54	1.92	0.13	-1.54
HP-P-OK1-CAP	-1.75	1.36	-3.73	1.30	2.02	0.13	-1.53
HP-P-OK2	-0.02	1.08	-1.99	2.43	2.43	0.10	-2.82
HP-P-OK2-CAP	-0.01	1.05	-1.73	2.45	2.44	0.11	-2.84***

a: data in the columns (1)- (6) are in % points. Growth rates have been annualised in the column (5).
b: calculated as the growth rate of the trend output at the fourth quarter of the year 2000.
c: ADF Test. Test specification: no constant, no deterministic trend, four lags included.
***, ** , * denote rejection at the 1 (5,10) percent level.

Table 1.c: Estimated weights

Models	HP	HP-P	HP-P-CAP	HP-P-OK1	HP-P-OK1-CAP	HP-P-OK2	HP-P-OK2-CAP
Inflation equation	-	2	5	12	10	2	3
Okun's law equation	-	-	-	79	53	38	38
Capacity rate equation	-	-	3	-	101		0.2

N.B.: Estimation period: 1980:1 2000.4. Weights are calculated for each relationship as the ratio of the variance of the growth rate of the trend and the variance of the economic relationship. For the inflation equation weight, to allow a comparison with the literature, weights have been computed for year on year inflation rates.

Table 1.d: Values of the weights in the HPMV literature

Empirical study	Economic relations weights and smoothing parameters	
	$\sigma^2_{gap}/\sigma^2_{\Delta\Delta y^*}$	Economic relationships weights
Laxton and Tetlow 1992	1600	1
Butler 1996	1600 16 000 for the participation rate 10 000 for the labour-output elasticity (univariate HP filter)	1
Haltmaier 1996	1600	25 (values between 12-50 tested)
Conway and Hunt 1997	1600	2 (alternative values tested)
De Brouwer 1998	1600	Relative variances
Côté and Hostland 1993	Values between 10 and 100 000 tested 500 for output 100 for unemployment	Not clear

N.B: Weights of the inflation equation in HPMV filters are presented in this table, using year on year inflation rates. As Haltmaier (1996) uses quarter on quarter inflation rates and, hence, considers for the inflation equation weights between 200 and 800, these weights have been re-computed here on the basis of year on year inflation rates.

Assessment of the output gap estimates

Table 2.a: Uncertainty statistics (1981:1 2000:4)

Models	HP	HP-P	HP-P-CAP	HP-P-OK1	HP-P-OK1-CAP	HP-P-OK2	HP-P-OK2-CAP
Standard deviation of the output gap							
Smoothed output gap	0.83	0.84	1.14	1.46	1.36	1.08	1.07
Filtered output gap	0.94	0.95	1.40	1.54	1.43	0.98	1.02
Standard errors							
Smoothed output gap							
Total	0.28	0.28	0.24	0.33	0.32	0.16	0.18
Filter	0.28	0.28	0.22	0.16	0.16	0.16	0.16
Parameters	0.00	0.06	0.09	0.27	0.27	0.05	0.09
Filtered output gap							
Total	0.54	0.56	0.27	0.52	0.50	0.37	0.38
Filter	0.54	0.55	0.22	0.40	0.40	0.36	0.36
Parameters	0.00	0.10	0.15	0.29	0.28	0.08	0.12
Standard revisions							
Smoothed versus filtered output gaps	1.09	1.04	0.71	0.44	0.43	0.48	0.47

Table 2.b: Final and quasi real time output gap estimates (1997:1 2000:3)

Models	MEAN	SD	MIN	MAX	COR
HP					
Final (smoothed)	-0.01	0.51	-0.89	0.88	1
Quasi real time	0.75	0.34	0.04	1.2	0.88
HP-P					
Final	0.01	0.54	-0.91	0.95	1
Quasi real time	0.74	0.36	0.01	1.32	0.91
HP-CAP					
Final	-0.16	0.50	-1.07	0.68	1
Quasi real time	0.04	0.54	-1.26	0.72	0.94
HP-OK2					
Final	0.14	0.86	-1.30	1.62	1
Quasi real time	1.70	0.50	0.70	2.51	0.84
HP-OK2-CAP					
Final	0.09	0.80	-1.27	1.47	1
Quasi real time	1.29	0.65	-0.20	2.51	0.82

N.B.: MEAN: the mean of the output gap, SD: the standard deviation, MIN: the minimum value, MAX: the maximum value, COR: the correlation with the final estimate of the output gap.

Table 2.c : Revision statistics (final versus quasi real time output gaps (1997:1 2000:3))

Model	MEAN	SD	NS
HP	-0.76	0.26	0.51
HP-P	-0.72	0.26	0.47
HP-P-CAP	-0.20	0.18	0.35
HP-P-OK2	-1.57	0.51	0.59
HP-P-OK2-CAP	-1.20	0.45	0.57

N.B.: MEAN : mean of the revisions, SD: Standard deviation of the revisions, NS: ratio of the standard deviation of the revision and the standard deviation of the final estimate of the output gap.

Table 2.d: Inflation forecast

Model	Naive model : AR(2) model				Naive model : random walk			
	RMSE	RRMSE	D-M Stat.	P value	RMSE	RRMSE	D-M Stat.	P value
HP-P	0.187	0.937	-0.789	0.428	0.187	0.799	-2.769	0.006
HP-P-CAP	0.181	0.908	-1.331	0.184	0.181	0.775	-3.074	0.001
HP-P-OK1	0.179	0.899	-1.543	0.122	0.179	0.767	-3.044	0.001
HP-P-OK2	0.185	0.927	-0.885	0.376	0.185	0.792	-2.852	0.002
HP-P-OK1-CAP	0.179	0.899	-1.562	0.118	0.179	0.767	-3.056	0.001
HP-P-OK2-CAP	0.184	0.923	-0.950	0.342	0.184	0.787	-2.912	0.001
Naive	0.199	1.000	0.000	1.000	0.234	1.000	0.000	1.000

N.B. The table compares the one step ahead inflation forecast (conditional to the smoothed state series) of the alternative models to the one step ahead forecast of two naive models (an AR(2) model and a random walk). The RMSE (root mean squared error) is the square root of the mean prediction errors. The RRMSE isthe relative root mean squared error of the models against that of the naive models. The D-M statistics is the Diebold-Mariano (Diebold and Mariano 1995) test statistics (see Appendix 6.4). P value is the marginal probability of the Diebold-Mariano test of the null hypothesis of equal predictive accuracy of the alternative models and of the naive model.

Impact of varying weights

Table 3.a: Parameter estimates of the economic relationships

	Model HP-P (weight=2)	Model HP-P (weight=5)	Model HP-P (weight=25)
π_{t-1}	0.33	0.31	0.25
	(0.00)	(0.00)	(0.00)
π_{t-1}	0.25	0.25	0.25
	(0.07)	(0.01)	(0.00)
$Y_{t-1}-Y^*_{t-1}$	0.03	0.04	0.06
	(0.28)	(0.00)	(0.00)
S_{t-1}	0.08	0.09	0.09
	(0.00)	(0.00)	(0.00)
S_{t-2}	-0.05	-0.05	-0.06
	(0.03)	(0.00)	(0.00)
σ^2_π	3.54e-06	2.35e-06	1.78e-06
	(0.00)		
HP parameter			
$\sigma^2_{\Delta\Delta y^*}$	7.44 e-08	1.47e-07	4.45e-07
	(0.00)	(0.00)	(0.00)

N.B.: Estimation period: 1980:1 2000.4

Table 3.b: Properties of the HPMV output gaps for alternative weights

	SD	Standard errors			Standard revision (smoothed versus filtered output gap)	Inflation forecast RRMSE	
		Total	Filter	Parameter		AR(2)	RW
HP-P (weight = 2)	0.84	0.28	0.28	0.06	1.04	0.94	0.80
HP-P (weight = 5)	0.89	0.37	0.37	0.08	0.96	0.92	0.78
HP-P (weight = 25)	1.13	0.50	0.48	0.14	0.99	0.87	0.74

Graphs of the HPMV output gap estimates

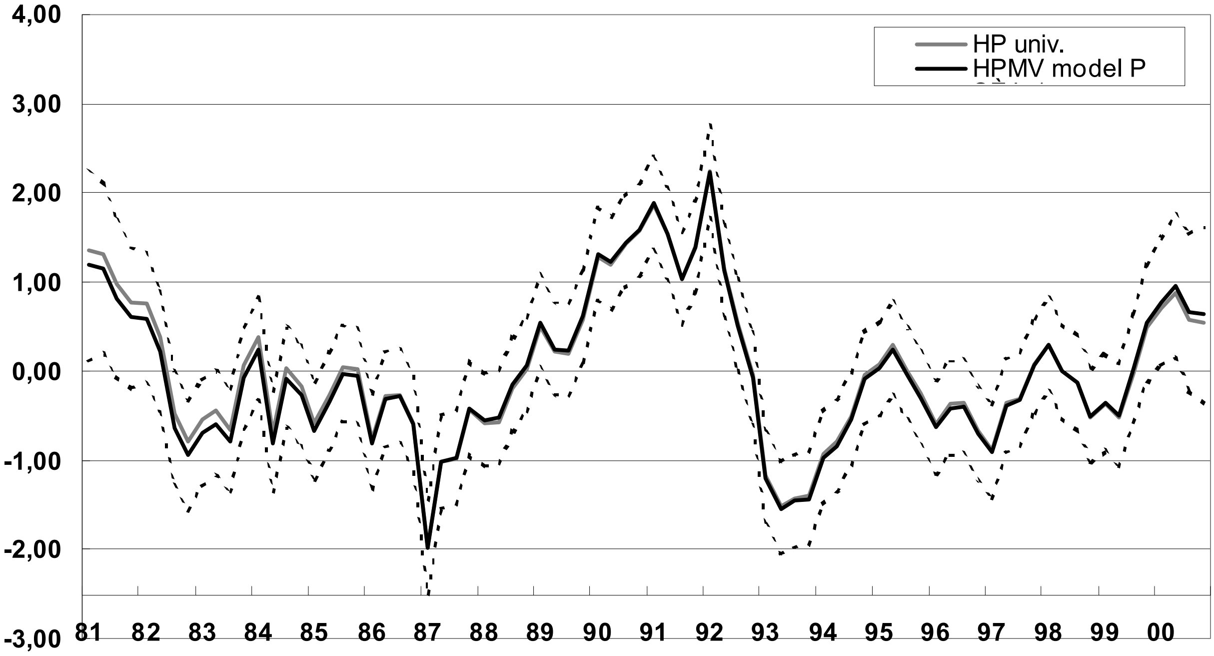

Figure 1: HP-P model

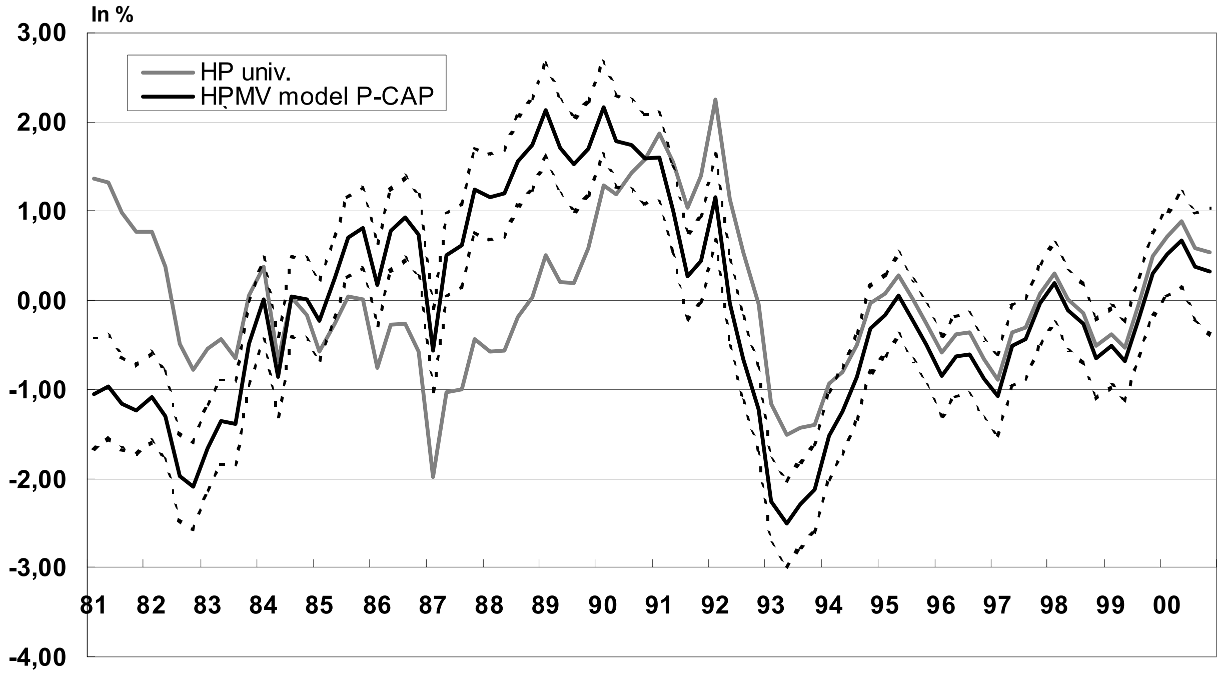

Figure 2: HP-P-CAP model

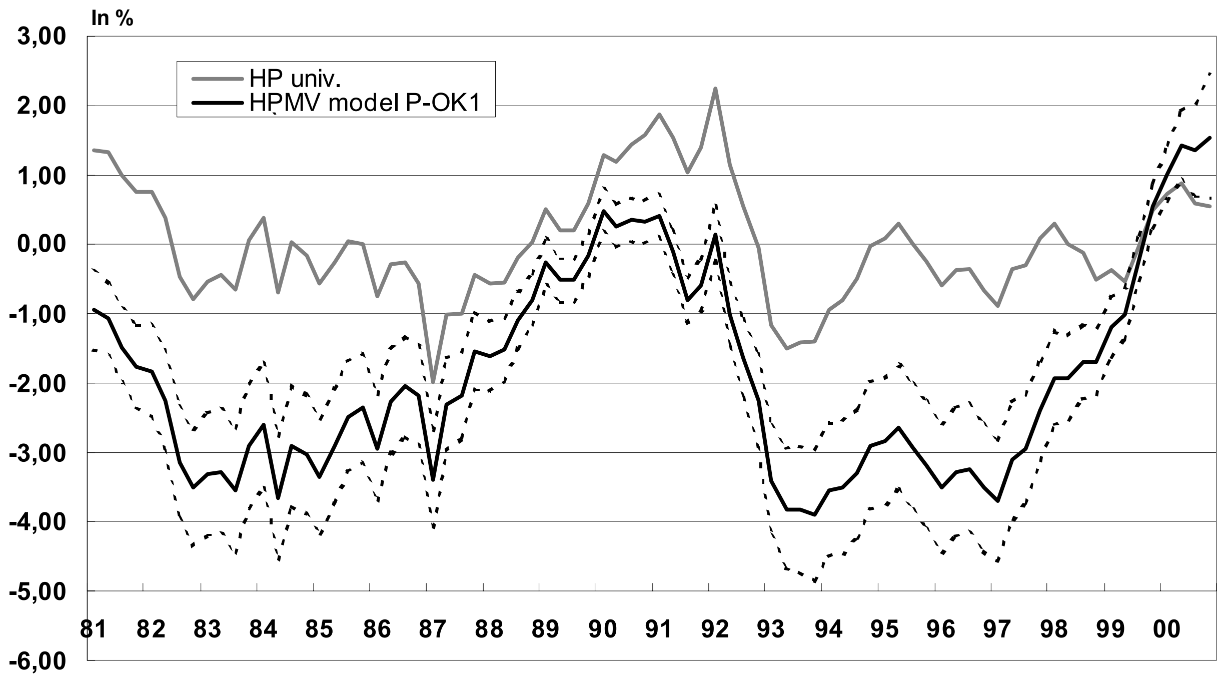

Figure 3: HP-P-OK1 model

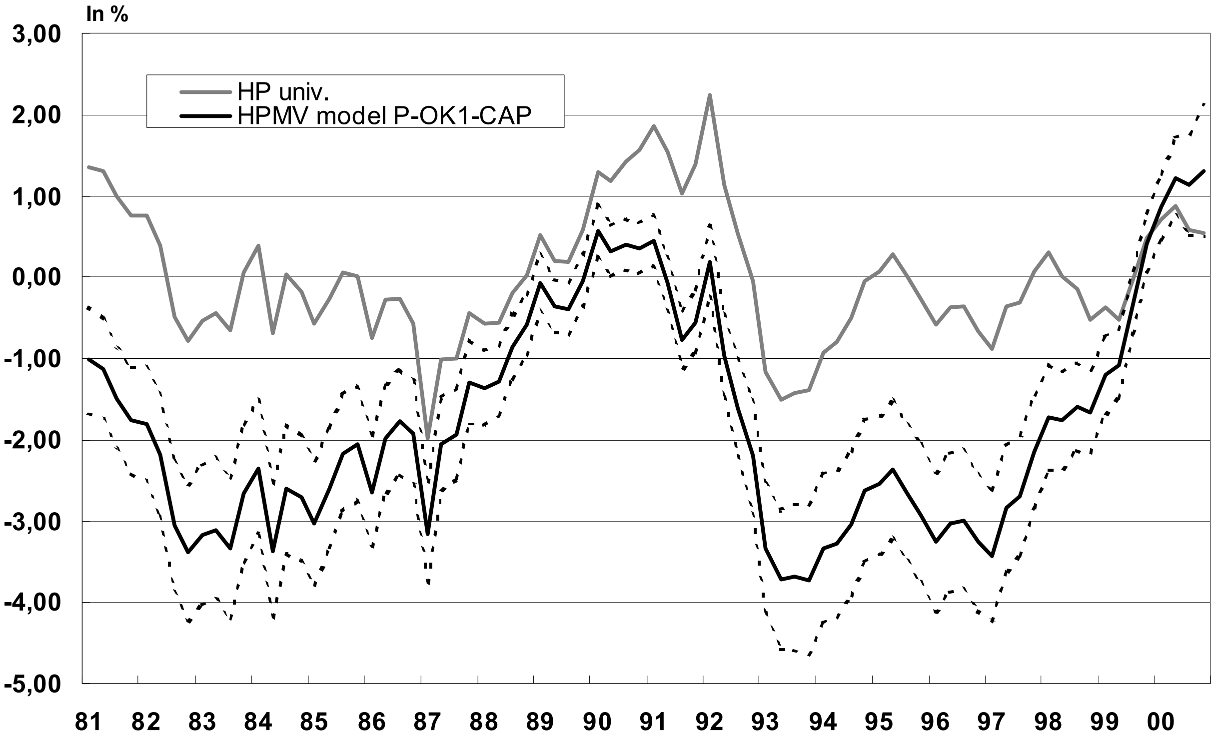

Figure 4: HP-P-OK1-CAP model

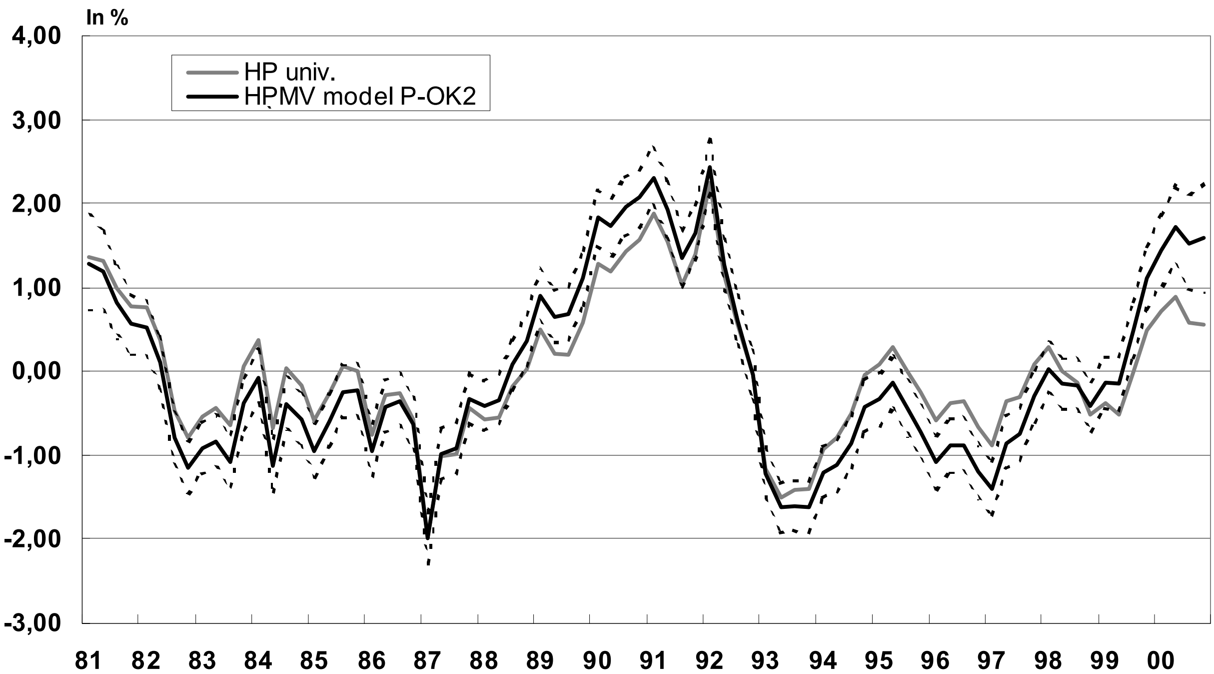

Figure 5: HP-P-OK2 model

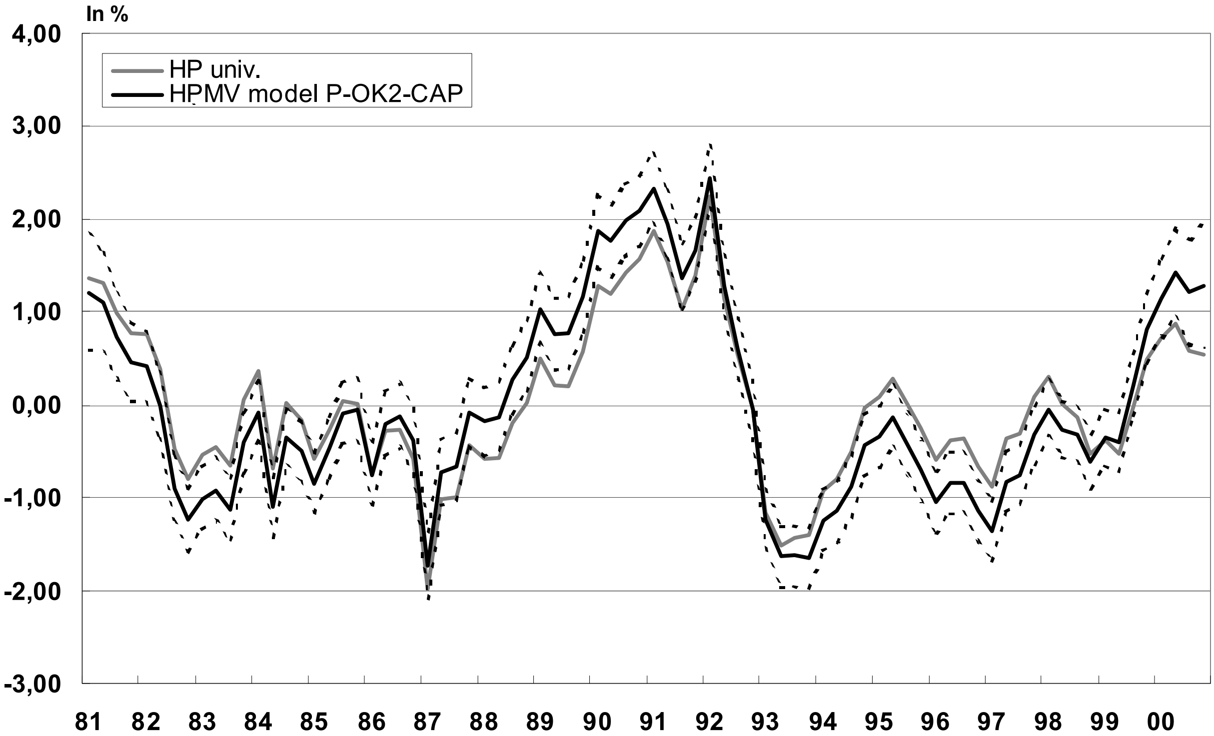

Figure 6: HP-P-OK2-CAP model

Figure 7: Output gap of the Euro area (HP-P model) using alternative weights for the inflation equation Phillips curve

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Detrending Techniques"

MODELLING CORE INFLATION FOR THE UK USING A NEW DYNAMIC FACTOR ESTIMATION METHOD AND A LARGE DISAGREGATED PRICE INDEX DATASET

By George Kapetanios[1]

Queen Mary, University of London

[1]Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS. E-mail: G.Kapetanios@qmul.ac.uk

1 Introduction

Recent work in the macroeconometric literature considers the problem of summarising efficiently a large set of variables and using this summary for a variety of purposes including forecasting.

Work in this field has been carried out in a series of recent papers by Stock and Watson (1998), Forni and Reichlin (2000), Forni, Hallin, Lippi, and Reichlin (2000) and Forni, Hallin, Lippi, and Reichlin (2001). Factor analysis has been the main tool used in summarising the large datasets.

The main factor model used in the past to extract dynamic factors from economic time series has been a state space model estimated using maximum likelihood. This model was used in conjunction with the Kalman filter in a number of papers carrying out factor analysis (see, among others, Stock and Watson (1989) and Camba-Mendez, Kapetanios, Smith, and Weale (2001)). However, maximum likelihood estimation of a state space model is not practical when the dimension of the model becomes too large due to the computational cost. For the case considered by Stock and Watson (1998) where the number of time series is greater than the number of observations, maximum likelihood estimation is not practically feasible. For this reason, Stock and Watson (1998) have suggested an approximate dynamic factor model based on principal component analysis.

This model can accommodate a very large number of time series and there is no need for the number of observations to exceed the number of variables. Nevertheless, the principal component model is not, strictly speaking, a dynamic model. Stock and Watson (1998) have shown that it can estimate consistently the factor space asymptotically (but the number of time series has to tend to infinity). In small samples and for a finite number of series, the dynamic element of the principal component analysis is not easy to interpret. Forni and Reichlin (2000) suggested an alternative procedure based on dynamic principal components (see Brillinger (1981, ch. 9)). This method incorporates an explicitly dynamic element in the construction of the factors.

This paper discusses an alternative method for estimating factors derived from a factor state space model. This model has a clear dynamic interpretation. Further, the method does not require iterative estimation techniques and due to a modification introduced, can accommodate cases where the number of variables exceeds the number of observations.

The computational cost and robustness of the method is comparable to that of principal component analysis because matrix algebraic methods are used. The method forms parts of a large set of algorithms used in the engineering literature for estimating state space models called subspace algorithms. Another advantage of the method is that the asymptotic distribution and therefore the standard errors of the factor estimates are available. Further, as the factor analysis is carried out within a general model, forecasting is easier to carry out than in the currently available procedures where a forecasting model needs to be specified.

The structure of the paper is as follows: Section 2 describes the elements of the suggested factor extraction method. Sections 3-5 discuss aspects of the new methodology. Section 6 presents an application of the method to the extraction of core inflation and forecasting of UK inflation in the recent past. Section 7 concludes.

2 The method

We consider the following state space model[2].

$$
\begin{aligned}
x_t &= Cf_t + Du_t, \quad t = 1, \dots, T \\
f_t &= Af_{t-1} + Bu_{t-1}
\end{aligned}
\tag{2}
$$

x_t is an n-dimensional vector of strictly stationary zero-mean variables observed at time t. f_t is an m-dimensional vector of unobserved states (factors) at time t and u_t is a multivariate standard white noise sequence of dimension n. The aim of the analysis is to obtain estimates of the states f_t, for $t = 1, \dots, T$.

This model is quite general. Its aim is to use the states as a summary of the information available from the past on the future evolution of the system. A large literature exists on the identification issues related with the state space representation given in (2). An extensive discussion may be found in Hannan and Deistler (1988). As we have mentioned in the introduction, maximum likelihood techniques either using the Kalman filter or otherwise may be used to estimate the parameters of the model under some identification scheme. For large datasets this is likely to be computationally intensive. Subspace algorithms avoid expensive iterative techniques and instead rely on matrix algebraic methods to provide estimates for the factors as well as the parameters of the state space representation.

There are many subspace algorithms and vary in many respects but a unifying characteristic is their view of the state as the interface between the past and the future in the sense that the best linear prediction of the future of the observed series is a linear function of the state. A very good review of existing subspace algorithms is given by Bauer (1998) in an econometric context. Another review with an engineering perspective may be found in Van Overschee and De Moor (1996).

The starting point of most subspace algorithms is the following representation of the system which follows from the state space representation and the assumed nonsingularity of D.

$$
X_t^f = \mathcal{O}\mathcal{K}X_t^p + \mathcal{E}E_t^f \tag{3}
$$

where $X_t^f = (x_t', x_{t+1}', x_{t+2}', \dots)'$, $X_t^p = (x_{t-1}', x_{t-2}', x_{t-3}', \dots)'$, $E_t^f = (u_t', u_{t+1}', u_{t+2}', \dots)'$, $\mathcal{O} = [C', A'C', (A^2)'C', \dots]'$, $\mathcal{K} = [\bar{B}, (A - \bar{B}C)\bar{B}, (A - \bar{B}C)^2\bar{B}, \dots]$, $\bar{B} = BD^{-1}$ and

$$
\mathcal{E} = \begin{pmatrix} D & 0 & \dots & 0 \\ CB & D & \ddots & \vdots \\ CAB & \ddots & \ddots & 0 \\ \vdots & & CB & D \end{pmatrix}
$$

The derivation of this representation is easy to see once we note that (i) $X_t^f = \mathcal{O}f_t + \mathcal{E}E_t^f$ and (ii) $f_t = \mathcal{K}X_t^p$. The best linear predictor of the future of the series at time t is given by

[2] Note that the model we present is equivalent to the more common form given by

$$
\begin{aligned}
x_t &= Cf_t + u_t, \quad t = 1, \dots, T \\
f_t &= Af_{t-1} + v_t
\end{aligned}
\tag{1}
$$

as proven in Hannan and Deistler (1988, pp. 17-18).

$\mathcal{OK}X_t^p$. The state is given in this context by $\mathcal{K}X_t^p$ at time t. The task is therefore to provide an estimate for $\mathcal{K}$. Obviously, the above representation involves infinite dimensional vectors.

In practice, truncation is used to end up with finite sample approximations given by $X_{s,t}^f = (x_t', x_{t+1}', x_{t+2}', \ldots, x_{t+s-1}')'$ and $X_{q,t}^p = (x_{t-1}', x_{t-2}', \ldots, x_{t-q}')'$. Then an estimate of $\mathcal{F} = \mathcal{OK}$ may be obtained by regressing $X_{s,t}^f$ on $X_{q,t}^p$. Following that, the most popular subspace algorithms use a singular value decomposition of an appropriately weighted version of the least squares estimate of $\mathcal{F}$, denoted by $\hat{\mathcal{F}}$. In particular the algorithm we will use, due to Larimore (1983), applies a singular value decomposition to $\hat{\Gamma}^f \hat{\mathcal{F}} \hat{\Gamma}^p$, where $\hat{\Gamma}^f$, and $\hat{\Gamma}^p$ are the sample covariances of $X_{s,t}^f$ and $X_{q,t}^p$ respectively. These weights are used to determine the importance of certain directions in $\hat{\mathcal{F}}$. Then, the estimate of $\mathcal{K}$ is given by

$$\hat{\mathcal{K}} = \hat{S}_m^{1/2} \hat{V}_m \hat{\Gamma}^{p^{-1}}$$

where $\hat{U}\hat{S}\hat{V}$ represents the singular value decomposition of $\hat{\Gamma}^f \hat{\mathcal{F}} \hat{\Gamma}^p$, $\hat{S}_m$ denotes the matrix containing the first m columns of $\hat{S}$ and $\hat{V}_m$ denotes the heading $m \times m$ submatrix of $\hat{V}$. $\hat{S}$ contains the singular values of $\hat{\Gamma}^f \hat{\mathcal{F}} \hat{\Gamma}^p$ in decreasing order. Then, the factor estimates are given by $\hat{\mathcal{K}}X_t^p$. For what follows it is important to note that the choice of the weighting matrices are important but not crucial for the asymptotic properties of the estimation method. They are only required to be nonsingular. A second thing to note is that consistent estimation of the factor space requires that q tends to infinity at a certain rate as T tends to infinity as pointed out by Bauer (1998, pp. 54). Once estimates of the factors have been obtained and if estimates of the parameters (including the factor loadings) are subsequently required, it is easy to see that least squares methods may be used to obtain such estimates. These estimates have been proved to be $\sqrt{T}$-consistent and asymptotically normal in Bauer (1998, ch.4). We note that the identification scheme used above is implicit and depends on the normalisation used in the computation of the singular value decomposition. Finally, we must note that the method is also applicable in the case of unbalanced panels. In analogy to the work of Stock and Watson (1998) use of the EM algorithm, described there, can be made to provide estimates both of the factors and of the missing elements in the dataset.

3 Dealing with large datasets

Up to now we have outlined an existing method for estimating factors which requires that the number of observations be larger than the number of elements in X_t^p. Given the work of Stock and Watson (1998) this is rather restrictive. We therefore suggest a modification of the existing methodology to allow the number of series in X_t^p be larger than the number of observations. The problem arises in this method because the least squares estimate of $\mathcal{F}$ does not exists due to rank deficiency of $X^{p'}X^p$ where $X^p = (X_1^p, \ldots, X_T^p)'$. As we mentioned in the previous section we do not necessarily want an estimate of $\mathcal{F}$ but an estimate of the states $X^p\mathcal{K}'$. That could be obtained if we had an estimate of $X^p\mathcal{F}'$ and used a singular value decomposition of that. But it is well known (see e.g. Magnus and Neudecker (1988)) that although $\hat{\mathcal{F}}$ may not be estimable $X^p\mathcal{F}'$ always is using least squares methods. In particular, the least squares estimate of $X^p\mathcal{F}'$ is given by

$$\widehat{X^p\mathcal{F}'} = X^p(X^{p'}X^p)^+ X^{p'}X^f$$

where $X^f = (X_1^f, \ldots, X_T^f)'$ and A^+ denotes the unique Moore-Penrose inverse of matrix A. Once this step is modified then the estimate of the factors may be straightforwardly obtained by applying a singular value decomposition to $\widehat{X^p\mathcal{F}'}$. We choose to set both weighting matrices to the identity matrix in this case.

4 Number of factors

A very important question relates to the determination of the number of factors, i.e. the dimension of the state vector. This issue has only recently received attention in the econometric literature. Stock and Watson (1998) suggest using information criteria for determining this dimension. Bai and Ng (2002) provide modified information criteria and justification for their use in the case where the number of variables goes to infinity as well as the number of observations. We suggest a simple information theoretic method for determining the number of factors in our model. Its simplicity comes from the fact that both the number of series and factors are assumed to be finite.

The search simply involves (i) fixing a maximum number of factors f^{max} to search over, (ii) estimating the factors for each assumed number of factors $m = 1, \ldots, m^{max}$ and (iii) minimising the negative penalised loglikelhood of the regression

$$x_t = C\hat{f}_t + u_t,$$

i.e. minimising $ln|\hat{\Sigma}_u^m| + c_T(m)$ where $\hat{\Sigma}_u^m$ is the estimated covariance matrix of u_t and $c_T(m)$ is a penalty term depending on the choice of the information criterion used. The theoretical properties of the new methodology are discussed in detail in Kapetanios (2002).
We briefly discuss an alternative class of testing procedures for determining the number of factors prevalent in the state space model literature. The testing procedures are based on the well known fact that the rank of certain block matrices referred to as Hankel matrices is equal to the dimension of the state vector. The most familiar Hankel matrix is the covariance Hankel matrix. The autocovariance Hankel matrix is a block matrix made up of the autocovariances of the observed process x_t. It is given by

$$\begin{pmatrix} \Gamma_1 & \Gamma_2 & \Gamma_3 \ldots \\ \Gamma_2 & \Gamma_3 & \ldots \\ \Gamma_3 & \ldots & \ldots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

where Γ_i denotes the i-th autocovariance of x_t. Its finite truncation may be estimated by $1/TX^{f'}X^p$. Tests of rank may be used to estimate the rank of the covariance Hankel matrix from its estimate. A thorough investigation of the properties of the information criteria and the testing procedures in determining the rank of the Hankel matrix may be found in Camba-Mendez and Kapetanios (2001b). Further issues are discussed in Camba-Mendez and Kapetanios (2001a). A related discussion of the tests of rank used may also be found in Camba-Mendez, Kapetanios, Smith, and Weale (2000).

5 Extensions

The analysis of large datasets based on a state space model and estimated using subspace methods can be extended in a number of ways. Up to now we have not entertained the possibility of idiosyncratic serially correlated errors for particular variables. This extension is straightforward in the state space model context, as these errors may simply be modelled as extra factors, that enter one or a few variables. In that sense the analysis does not change. However, one may wish to draw a more clear distinction between common factors and idiosyncratic errors. Such a distinction can be accommodated by assuming that the number of variables tends to infinity following the ideas of Stock and Watson (1998). Crucially, the computational aspects of the analysis do not change.

Another important extension can be envisaged in terms of developing structural models for large datasets in the spirit of structural VAR (SVAR) models popularised in the 90's. Considering the state space model of the form

$$\begin{aligned} x_t &= Cf_t + u_t, \quad t = 1, \ldots, T \\ f_t &= Af_{t-1} + v_t \end{aligned} \tag{4}$$

we may distinguish between the shocks u_t and v_t and attribute structural meaning to linear combinations of v_t following the SVAR literature. Many possible identification schemes are possible and research in them is carried out in Kapetanios and Marcellino (2002).

6 An Application: Extracting Core Inflation

In this section we provide an application of the dynamic factor methodology to the modelling of UK core inflation. We take as our measure of inflation the RPIX (RPI minus mortgage interest payments) inflation used by the Bank of England at the target measure for monetary policy. Core inflation is a fuzzy concept which has been defined in various ways in the literature. We will not attempt to provide even a partial review of a huge literature. In general, when people use the term core inflation they seem to refer to the long-run or persistent component of the measured price index. A clear definition of core inflation requires a model of how prices and money are determined in the economy. We choose to follow an atheoretical approach to the definition of core inflation by specifying it to be the major dynamic factor underlying the components used to construct the retail price index.

More specifically let the set of individual price component growth rates be denoted by x_t. These growth rates are obtained by differencing the logarithm of the respective component price index. Then, x_t is specified to follow a model of the form (4). Core inflation at time t is defined to be the first factor in the vector f_t as defined by the ordered singular values of the singular value decomposition of $\mathcal{F} = \mathcal{O}\mathcal{K}$ in (3). This definition although in no way related to a theoretical economic model is consistent with the prior idea that core inflation is the main persistent component of inflation.

We fit a state space model to the components of the RPIX price index for the period of January 1987 to August 2002. Monthly data are used. Information on the components used are given in the data appendix. We set the truncation indices to $s = 1$ and $q = 3$ respectively. We note that q has to tend to infinity as the sample size grows in order to get a consistent estimate of the factors. We have chosen to set this to 3 because the resulting estimate of core inflation does not change perceptibly as q is increased from this value. Component series were normalised to have mean equal to zero and variance equal to one prior to estimation of the factor. We present RPIX inflation and our measure of the core inflation in Figure 1.

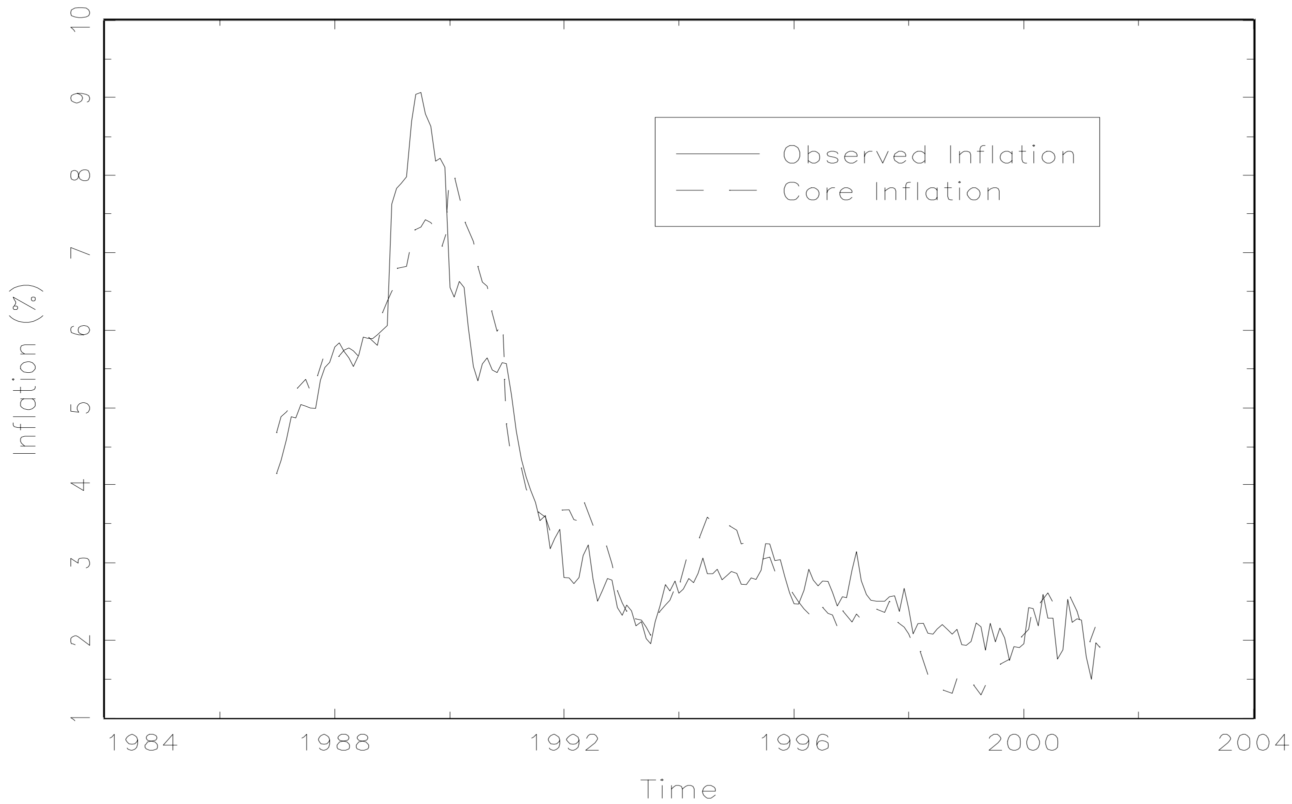

Figure 1: Observed and core inflation

Note that the core inflation has been normalised to have the same mean and variance as observed inflation over the sample period. Clearly, the factor model estimate of the core inflation is smoother than actual observed inflation. However, at business cycle frequencies it exhibits pronounced cyclicality. The departure from observed inflation in the spike of the late eighties and early nineties can be traced back to tax changes (including the repeal of the poll tax) in that period. Our measure of core inflation can explain on average 44% of a given component series whereas addition of an extra factor raises this to 53%.

Having obtained a means of estimating core inflation we now examine the forecasting abilities of this measure. In particular we consider three models. One is a simple benchmark AR model where the lag order is chosen automatically using the Akaike information criterion. The second is the benchmark model augmented by the growth rate of money and in particular M0. Lag selection is again carried out by the Akaike information criterion for both inflation and the money growth rate. Finally, the third model is the benchmark model augmented with the currently available estimate of the core inflation.

We evaluate the three models over the period June 1998-August 2002. We have allowed for a year following the introduction of independence for the Bank of England to carry out monetary policy though an inflation targeting regime. We examine both relative RMSEs compared to the model which includes the factor and the Diebold and Mariano (1995) test statistic for equality in predictive ability between two different forecasts. All models are estimated recursively (including lag order selection). The forecasts are examined for horizons of 1 to 4 months ahead. All results are presented in Table 1.

Table 1: Results on forecasting performance

Horizon	DM [a]	DM [b]	RMSE [c]	RMSE [d]
1	1.42	0.66	0.95	0.68
2	0.13	0.26	0.99	0.98
3	0.48	0.67	0.97	0.92
4	0.61	1.04	0.97	0.89

[a] Diebold-Mariano test statistics against benchmark AR model
[b] Diebold-Mariano test statistics against money growth rate model
[c] Relative RMSE compared to benchmark AR model. Values less than 1 indicate superiority of factor model
[d] Relative RMSE compared to money growth rate model. Values less than 1 indicate superiority of factor model

The results show that the factor model can indeed help in forecasting. The factor model performs 32% better than the money growth model for forecasts one month ahead. The factor model always has a lower RMSE compared to the other models. Although the factor model may appear to have a similar performance compared to the AR model the Diebold-Mariano statistic, although not rejecting in favour of the factor model, indicates that with a probability value of 0.078 is close to rejection.

7 Conclusion

In this paper we have discussed a new factor based method for forecasting time series introduced by Kapetanios (2002). This work follows closely in spirit the work of Stock and Watson (1998), Stock and Watson (1999) and subsequent, as yet unpublished papers by these authors and their co-authors on the one hand and the work by Forni and Reichlin (2000), Forni, Hallin, Lippi, and Reichlin (2000) and Forni, Hallin, Lippi, and Reichlin (2001) on the other hand. The innovation lies in providing an alternative method for obtaining factor estimates.

One strand of the literature on factor extraction relies on explicitly dynamic state space models to estimate factors via computationally expensive and, in small samples, non-robust maximum likelihood estimation. The other strand of the literature based on the work of Stock and Watson (1998) uses principal components to extract the factors. This methodology is robust, computationally feasible with very large datasets and asymptotically valid for dynamic settings. Unfortunately, these methods are approximately dynamic in that the dynamic structure of the factors is not explicitly modelled in finite samples but captured only asymptotically where both the number of observations and the number of series used, grows to infinity. We propose a new methodology which while sharing all the advantages of the principal component extraction method is explicitly dynamic. This method is based on linear algebraic techniques for estimating the state and, if need be, the parameters of a general linear state space model.

We evaluate the new methodology by investigating a model of core inflation for the UK. The measure of core inflation obtained is shown have predictive ability for inflation in the UK over a relatively long evaluation period.

References

Bai, J., and S. Ng (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1).

Bauer, D. (1998): "Some Asymptotic Theory for the Estimation of Linear Systems Using Maximum Likelihood Methods or Subspace Algorithms," Ph.d. Thesis.

Brillinger, D. (1981): *Time Series: Data Analysis and Theory.* Holden-Day, San Francisco.

Camba-Mendez, G., and G. Kapetanios (2001a): "Statistical Testing Procedures and Information Criteria in Rank Determination for System Identification," Mimeo.

——— (2001b): "Testing the Rank of the Hankel Covariance matrix: A Statistical Approach," *Institute of Electrical and Electronic Engineers Transactions on Automatic Control*, 46(2), 331–336.

Camba-Mendez, G., G. Kapetanios, R. J. Smith, and M. R. Weale (2000): "Tests of Rank in Reduced Rank Regression Models," Forthcoming in Journal of Business and Economic Statistics.

——— (2001): "An Automatic Leading Indicator of Economic Activity: Forecasting GDP growth for European Countries," *Econometrics Journal*, 4, 56–90.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): "The Generalized Factor Model: Identification and Estimation," *Review of Economics and Statistics*, 82(4), 540–554.

——— (2001): "The Generalised Dynamic Factor Model: One Sided-Estimation and Forecasting," Mimeo, Université Libre de Bruxelles.

Forni, M., and L. Reichlin (2000): "Let's Get Real: A Factor Analytic Approach to Business Cycle Dynamics," *Review of Economic Studies*, 65(3), 453–473.

Hannan, E. J., and M. Deistler (1988): *The Statistical Theory of Linear Systems.* John Wiley.

Kapetanios, G. (2002): "Factor analysis using subspace factor models: Some theoretical results and an application to UK inflation forecasting," *Working paper no. 466, Queen Mary University of London, http://www.econ.qmul.ac.uk/papers/docs/wp466.pdf.*

Kapetanios, G., and M. Marcellino (2002): "Structural Impulse Response Analysis of Dynamic Factor Models for Large Datasets," *Mimeo.*

Larimore, W. E. (1983): "System Identification, Reduced Order Filters and Modelling via Canonical Variate Analysis," *Proc. 1983 Amer. Control Conference*, 2, 445–451.

Magnus, J. R., and H. Neudecker (1988): *Matrix Differential Calculus with Applications to Statistics and Econometrics.* John Wiley.

Stock, J. H., and M. Watson (1989): "New Indexes (Sic) of Coincident and Leading Indicators," *NBER Macroeconomics annual*, 4, 351–394.

Stock, J. H., and M. W. Watson (1998): "Diffusion Indices," Forthcoming in Journal of Business and Economic Statistics.

——— (1999): "Forecasting Inflation," *Journal of Monetary Economics*, 44(2), 293–335.

Van Overschee, P., and B. De Moor (1996): *Subspace Identification for Linear Systems*. Kluwer Academic Publishers.

Data Appendix

RPIX components and their ONS (Office of National Statistics) codes.
bread DOAA.M
cereals DOAB.M
biscuits DOAC.M
beef DOAD.M
lamb DOAE.M
pork DOAG.M
bacon DOAH.M
poultry DOAI.M
other meat DOAJ.M
fish DOAK.M
butter DOAM.M
oil and fat DOAN.M
cheese DOAO.M
eggs DOAP.M
milk DOAQ.M
milk products DOAR.M
tea DOAS.M
coffee DOAT.M
soft drink DOAU.M
sugar DOAV.M
sweets chocolates DOAW.M
potatoes DOAX.M
vegetables DOAZ.M
other foods DOBD.M
restaurant meals DOBE.M
canteen meals DOBF.M
take aways DOBG.M
beer DOBH.M
wine DOBK.M
cigarettes DOBN.M
other tobacco DOBO.M
rent DOBP.M
council tax DOBR.M
water DOBS.M
repairs and maintenance DOBT.M
DIY DOBU.M
insurance and ground rent DOBV.M
coal DOBW.M
electricity DOBX.M
gas DOBY.M
oil and other fuel DOBZ.M
furniture DOCA.M
furnishings DOCB.M
appliances DOCC.M

other eqpt DOCD.M
consumables DOCE.M
pet care DOCF.M
postage DOCG.M
telephones DOCH.M
dom services DOCI.M
fees and subs DOCJ.M
clothing men DOCK.M
clothing women DOCL.M
clothing children DOCM.M
clothing other DOCN.M
footwear DOCO.M
personal articles DOCP.M
chemist goods DOCQ.M
personal services DOCR.M
purchase cars DOCS.M
maintenance cars DOCT.M
petrol and oil DOCU.M
tax and insurance DOCV.M
rail fares DOCW.M
bus and coach fares DOCX.M
other travel DOCY.M
audio visual DOCZ.M
CDs tapes DODA.M
toys and sports goods DODB.M
books and newspapers DODC.M
garden products DODD.M
tv licences DODE.M
entertainment and other recreation DODF.M

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on “Detrending Techniques”

ALTERNATIVE LINEAR AND NON-LINEAR DETRENDING TECHNIQUES: A COMPARATIVE ANALYSIS BASED ON EURO-ZONE DATA

By Torben Mark Pedersen[1]

Ministry of Economic and Business Affairs

It is shown how to compare the distortionary effect of alternative business cycle filters by measuring the distortionary effect of the filters. The measurements are based on the ad hoc assumption that the “true” business cycle filter is an ideal high-pass filter or an ideal band-pass filter for a band of frequencies which is determined after measuring the duration of growth cycles or the business cycle component.

Keywords: Optimal filtering; Detrending; Business cycles.

JEL classification: E32; C22.

[1] Correspondence: Torben Mark Pedersen, Ministry of Economic and Business Affairs, Slotsholmsgade 10-12, DK-1216 Copenhagen K. Denmark. Tel.: +45 33 92 41 68. E-mail: tmp@oem.dk

I would like to thank participants at the colloquium for helpful comments. I would also like to thank COMSOL A/S, Denmark, for a license to MATLAB r13.

1 Introduction

Modern general equilibrium business cycle research is aimed at explaining the fluctuations about a trend of output and its components as well as factors of production, following Lucas (1977) who defined "business cycles" as "*movements about trend in gross national product*". Lucas' definition of the business cycle corresponds to the older concept of *growth cycles* (or "deviation cycles") developed by Mintz (1969). There is an early American tradition for statistical empirical business cycle research dating back to the work of Mitchell (1913, 1927) and his associates which culminated in the monumental work "Measuring Business Cycles" by Burns and Mitchell (1946) who defined business cycles as fluctuations in the *level* of some measure of aggregate economic activity. The work of Burns and Mitchell has been continued by the NBER and they publish a set of widely accepted dates of turning points of the U.S. business cycle.

The crucial question in modern empirical business cycle research is how to define the trend and how to extract the trend component from the economic time series so that what is left is something that can be interpreted as the business cycle component. Hodrick and Prescott (1980, 1997) developed a statistical *definition* of the business cycle component (or growth cycle) and developed a now widely used method for summarizing the dynamics of empirical facts about growth cycles in the form of sample second moments summarizing the variation and serial correlation between economic variables.

Since the work of Hodrick and Prescott, a large number of ad hoc statistical methods have been developed for filtering a stochastic trend from macroeconomic time series. Unfortunately, the computed second moments of filtered time series depend critically on the filtering method, as documented by King and Rebelo (1993), Canova (1994, 1998, 1999) and many others. If the purpose is to "measure business cycles", as claimed by Hodrick and Prescott (1980, 1997) and Baxter and King (1999) among others, then we need a method for determining which set of stylized facts to trust. That is, we need, first, a measure of the true business cycle component, secondly, an estimate of the actual business cycle component after filtering with some filter and, finally, a metric for measuring the distance between the true and the estimated business cycle component. It was shown in Pedersen (2001) that there are no general results about the distortionary effect of filters since it depends both on the filter being used and the time series being filtered. It is the goal of this paper to compare the distortionary effect of alternative filters based on quarterly real GDP for the EU, the Euro-zone, Belgium, Denmark, Finland, France, Italy, Netherlands, Spain, Sweden, and the UK.

The plan of the paper is as follows. Filtering in the time and frequency domain is treated in Section 2. A metric for measuring the distortionary effect of filters is presented in Section 3, and a definition of the business cycle component in Section 4. Alternative filters are compared in Section 5. Summary and conclusions in Section 6.

2 Filtering

Let $\{y_t\}_{t=1}^{T}$ be a nonstationary, univariate time series such as logarithm of quarterly real GDP. When filtering macroeconomic time series, it is assumed that y_t can be decomposed into an unobservable trend or growth component, τ_t, an unobservable business cycle component, c_t, an unobservable seasonal component, s_t, and an unobservable irregular component ε_t[2]

$$y_t = \tau_t + c_t + s_t + \varepsilon_t, \tag{1}$$

where the different components are defined as cycles with different periods or as specific bands of frequencies. The unobserved components are assumed to be uncorrelated.

A zero correlation between the growth component and the cyclical component in (1) can be generated by a standard real business cycle model where stationary technology shocks to the production possibilities generate stationary fluctuations around a deterministic growth trend. In that class of models, shocks to the cyclical and secular components are orthogonal. That is not the case in endogenous growth models where shocks to the trend and the cyclical component are perfectly correlated. As should be obvious, the decomposition into a trend and a cyclical component only makes sense if growth is not endogenous.

Filtering is about removing something which is undesirable from the desirable as when filtering impurities from water or when filtering noise from a signal. Filtering is used widely in economics. Many signals or information sets are noisy and we need to extract the signal from the noise. Filtering is thus a way of solving a signal extraction problem. In Lucas (1972), rational agents solve a signal extraction problem in order to react optimally to an observed price change where it is unknown whether the price change reflects a change in the general price level or a change in real demand on the individual market.

To decompose the model (1) into its separate unobserved components is a signal extraction or filtering problem. Filtering problems can be analyzed both in the time and in the frequency domain. Where time is the fundamental unit in the time domain, the frequency is the fundamental unit in the frequency domain.

2.1 Filtering in the Time Domain

Let $\{x_t\}_{t=-\infty}^{\infty}$ be a stationary stochastic process with finite power (finite variance). Suppose that a new series $\{y_t\}_{t=-\infty}^{\infty}$ is constructed as a weighted moving average of lagged and future values of the input process $\{x_t\}_{t=-\infty}^{\infty}$

$$y_t = \sum_{j=-\infty}^{\infty} h_j x_{t-j} = h(L)x_t, \tag{2}$$

where the *filter weights* h_j are real and do not depend on time, $h(L) \equiv \sum_{j=-\infty}^{\infty} h_j L^j$ is a polynomial in the lag operator, L, with $\sum_{j=-\infty}^{\infty} |h_j| < \infty$, a requirement needed to assure that the variance of y_t is finite. We say that y_t is a filtered version of x_t.

When analyzing filters, we use the following terminology: The filter $h(L)$ is a *linear time-invariant filter* (or linear filter or LTI for short) since y_t is a linear function of x_t and the weights are independent of time. LTI filters take the form of weighted moving averages. The filter (2) is a two-sided filter because both past, present, and future values of x_t are used.

[2] Abstracting from a seasonal component if we study seasonally adjusted time series.

Two-sided filtering corresponds to the concept of *smoothing* as defined by Anderson and Moore (1979). We call it a *one-sided* or *backward-looking* filter if only past and present terms of x_t are involved. The filter is *symmetric* when $h_j = h_{-j}$, $j = 1, 2, ...$

Any LTI filter can be given a transfer function representation, a state space representation, a vector autoregressive representation, and a spectral representation so all LTI-filters with a spectral representation can be derived as an unobserved component model with a state space representation. The limitation of LTI-filters is exactly that they are limited to time-invariant filters with constant filter weights. This may seem like a critical assumption since business cycles and growth cycles are far from periodic and it could be argued that business cycles should be modelled as stochastic cycles.

One can think of y_t as the output process of an economic *model* where $\{x_t\}_{t=-\infty}^{\infty}$ is the stochastic shock process and the filter weights $h(L)$ are the economic model. Slutzky (1927, 1937) showed how filtering by successively summing and differencing a white noise process generates a time series with cycles - the so-called *Slutzky-effect* or Yule-Slutzky-effect, Yule (1927). The importance of that result for business cycle research lies in the fact that, if the economic model works like the summing and differencing operations, then exposing such a model to white noise shocks generates time series with cycles. Frisch (1933) shows that a model which consists of a system of difference and differential equations works like the summing and differencing filter of Slutzky and thus generates stochastic cycles in model generated time series of output, consumption, and investments when exposed to random shocks.

Real business cycle theory belongs to the Frisch-Slutzky approach to business cycle research where a stochastic dynamic general equilibrium model is exposed to an exogenous shock process and the economic model works like a filter which decomposes output into its components, consumption and investments.

2.2 Spectral Analysis and Filtering in the Frequency Domain

There is a general result known as the **spectral representation theorem** which says that any covariance-stationary stochastic process $\{y_t\}_{t=0}^{\infty}$ can be given a representation expressed in terms of an infinite weighted sum of (orthogonal) periodic functions of the form $\cos(\omega t)$ and $\sin(\omega t)$, where $\omega \in [-\pi, \pi]$ is the angular frequency measured in radians. Such a representation is called the **spectral representation** or Cramér's representation

$$y_t = \mu + \int_0^{\pi} \alpha(\omega) \cdot \cos(\omega t)\, d\omega + \int_0^{\pi} \delta(\omega) \cdot \sin(\omega t)\, d\omega. \qquad (3)$$

The goal of spectral analysis is to determine how important cycles of different frequencies are in accounting for the behavior of $\{y_t\}_{t=0}^{\infty}$.

Let $\{y_t\}_{t=-\infty}^{\infty}$ be a real-valued stationary stochastic process with variance equal to γ_0 and the τth autocovariance equal to γ_τ. Let $\{\gamma_\tau\}_{\tau=-\infty}^{\infty}$ be an absolutely summable sequence of autocovariances and let the autocovariance generating function in the frequency domain be given by: $g_y(z) \equiv \sum_{\tau=-\infty}^{\infty} \gamma_\tau z^\tau$, where the argument of the function, z, is a complex scalar.

The power spectrum (the power spectral density function) of $\{y_t\}_{t=-\infty}^{\infty}$ is given by

$$S_y(\omega) \equiv \frac{1}{2\pi} g_y\left(e^{-i\omega}\right) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau e^{-i\omega\tau}, \tag{4}$$

where $i \equiv \sqrt{-1}$.

It is seen from (4) that the *spectrum* of a stochastic process contains the same information as the autocovariance generating function since it is simply a linear combination of autocovariances.

The area under the spectrum over the interval $[-\pi, \pi]$ is equal to the variance of y_t and since the spectrum is symmetric around $\omega = 0$

$$var(y_t) \equiv \gamma_0 = 2 \int_0^\pi S_y(\omega)\, d\omega \tag{5}$$

and the τth autocovariance is

$$\gamma_\tau = 2 \int_{\omega=0}^{\pi} S_y(\omega)\, e^{i\omega\tau} d\omega. \tag{6}$$

This sheds some light on the interpretation of the variance in the time domain. The variance of a time series is generally (unless the process is white noise) distributed unevenly over frequencies where for example the growth component, the business cycle component, and the seasonal component contribute differently to the variance of y_t. The spectrum can be viewed as a device for decomposing the variance by frequency. If the business cycle component is defined as cycles between 5 quarters and six years, corresponding to the frequencies $\pi/12 \leq \omega \leq \pi/2.5$ with quarterly data, then the variance of the business cycle component (the growth cycle) is

$$var(c_t) = 2 \int_{\omega=\pi/12}^{\pi/2.5} S_y(\omega)\, d\omega. \tag{7}$$

A plot of the spectrum against frequencies shows how the variance of y_t is distributed over frequencies, Figure 1.[3] It is worth considering three cases:

1. If the estimated spectrum is flat without any peaks and without any clear tendency to follow a smooth curve, then the time series is close to a white noise.
2. If the time series contains a clear cyclical component at some frequency, then the spectrum has a tall, narrow peak (theoretically: infinitely tall, infinitely narrow peak, having finite area). Business cycles are not periodic and no sharp peaks at any frequencies should be expected.
3. A time series with an important trend component will have a strong peak at the very low frequencies, known as the "typical spectral shape", Granger (1966). If the time series contains an autoregressive unit root, then the peak will be infinite at zero frequency.

[3] I use the convention to plot the spectrum on the interval 0 to π because the spectrum is symmetric around $\omega = 0$.

2.2.1 Filtering in the Frequency Domain

It is possible to filter in the frequency domain by multiplying the spectrum of the input process at frequency ω by some non-negative number in order to get the spectrum of the output process at frequency ω. The filter in the frequency domain is an index with a non-negative real number at each frequency, and the index is called the *power transfer function*, denoted by $PTF(\omega)$.

Filtering a stationary stochastic process $\{y_t\}_{t=-\infty}^{\infty}$ with an LTI filter in the time domain,

$$c_t = C(L) \cdot y_t,$$

where $C(L)$ is a polynomial in the lag operator, has its parallel in the frequency domain where the spectrum of the output series at frequency ω is the power transfer function of the filter at frequency ω multiplied by the spectrum of the input series, $S_y(\omega)$ at frequency ω

$$S_c(\omega) = PTF(\omega) \cdot S_y(\omega), \tag{8}$$

where the power transfer function is the squared *frequency response function*

$$PTF(\omega) \equiv \left|h(e^{-i\omega})\right|^2$$

of the filter.

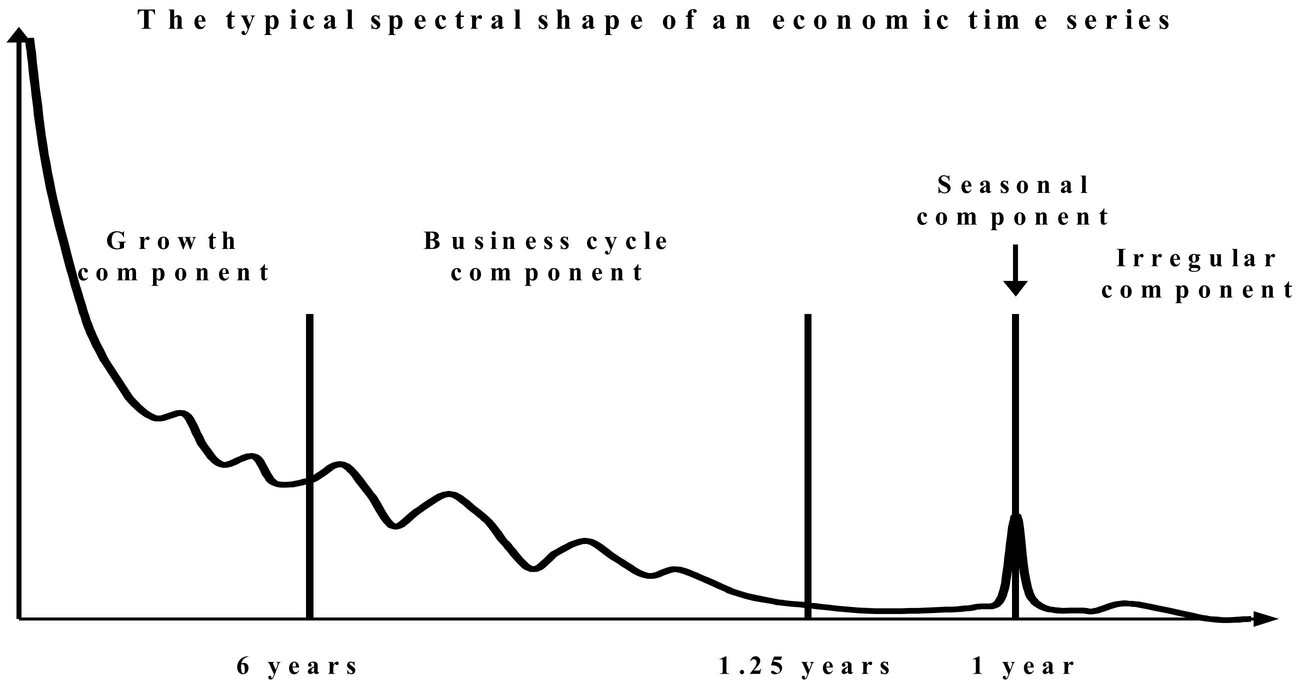

Figure 1: Decomposition of the spectrum of an economic time series with the "typical spectral shape" into bands of frequencies

2.3 The Effect of Filtering

The effect of linear filtering is twofold: To change the relative importance of the various cyclical components in (1) captured by the gain, and to induce a phase shift, Figure 2. Let x_t be a complex valued time series with the representation

$$x_t = \sum_{\omega} h_w e^{i\omega t}.$$

It can be given a polar decomposition as

$$x_t = \sum_{\omega} |h_w| \, e^{i(\omega t + \theta(\omega))},$$

where $|h_w|$ is the amplitude and $\theta(\omega)$ is the phase angle of h_ω. The frequency response function can be given a polar decomposition

$$h(e^{-i\omega}) = G(\omega) \cdot e^{iPh(\omega)}, \tag{9}$$

where the absolute value, $|G(\omega)|$, is the *gain* and $Ph(\omega)$ is the *phase*, then the effect of filtering x_t with a filter in the frequency domain is

$$h(\omega)x_t = \sum_{\omega} |h_w| \cdot G(\omega) \cdot e^{i(\omega t + \theta(\omega) + Ph(\omega))}. \tag{10}$$

That is, the amplitude at each frequency ω is multiplied by the gain of the filter and the phase is shifted from $\theta(\omega)$ to $\theta(\omega) + Ph(\omega)$, Figure 2. *Symmetric filters* have the attractive property that the phase shift is zero.

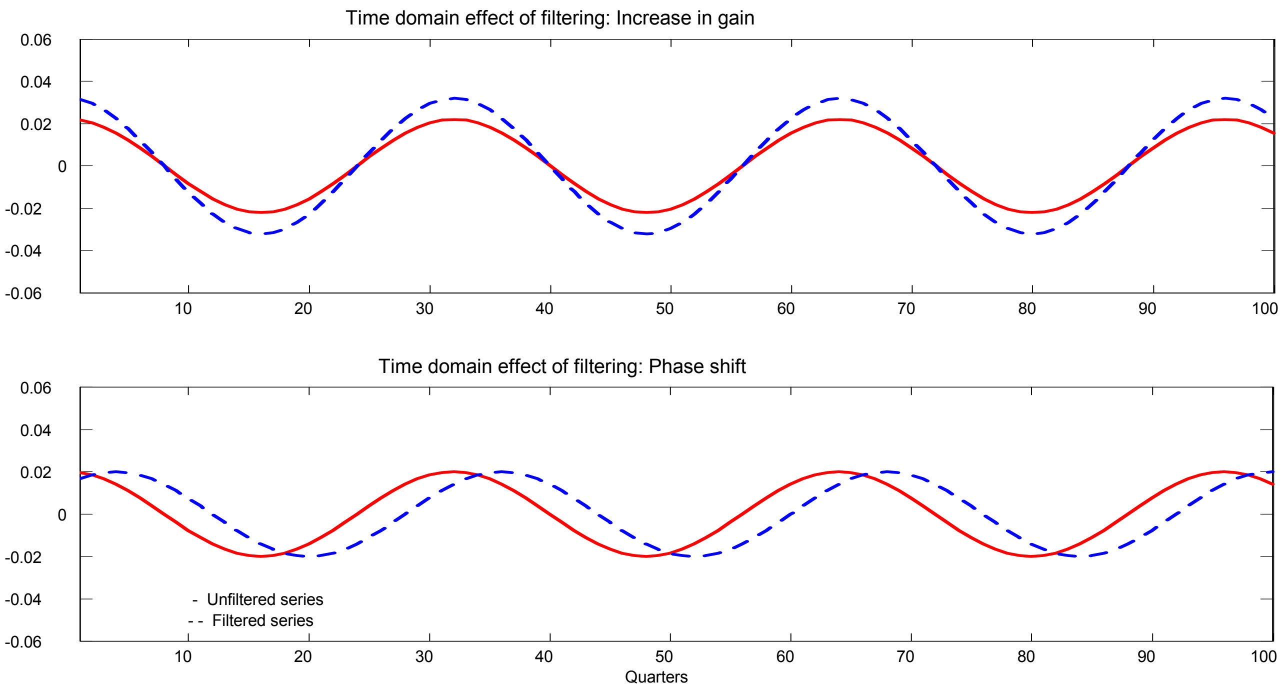

Figure 2: The effect of filtering: increase in gain and phase shift

The *gain* is defined as

$$G(\omega) = \frac{|S_{cy}(\omega)|}{S_{yy}(\omega)}, \tag{11}$$

which is the ratio between the cross-spectrum between c_t and y_t to the spectrum of y_t which is the regression coefficient at the ω-frequency of c_t on the ω-frequency of y_t, and so it can be interpreted as an expression of how the amplitude of y_t is multiplied in contributing to the amplitude in c_t. The gain of a time series measures - at the specified frequency ω - the increase in amplitude of one series over the other. The *phase* gives the lead of c_t over y_t at frequency ω. A *phase diagram* is the plot of the phase against frequency and it gives information about the lag relationship between two series, which series is leading or lagging at frequency ω. There is a close relationship between the concept of the phase and the business cycle research of identifying leading indicators.

The action of a filter is completely determined by the gain and phase, the frequency response function, or the power transfer function of the filter. A plot of the power transfer function against frequency shows which frequencies are attenuated and which are magnified.

2.4 Ideal Filters

A *low-pass filter*, denoted by lp, is a filter which allows low frequencies to pass the filter and removes all the variance above a prespecified cutoff frequency. A *high-pass filter* by contrast, hp, is one which allows high frequencies to pass the filter and removes all the variance at frequencies below a prespecified cutoff frequency. We can use the basic operations of combining linear filters to construct a high-pass filter from a low-pass filter: $hp(\omega) = 1 - lp(\omega)$. A *band-pass filter*, bp, is one which allows a specific band of frequencies to pass and removes higher and lower frequencies. A band-pass filter can be constructed as the difference between two low-pass filters: $bp(\omega) = lp_u(\omega) - lp_l(\omega)$, where $lp_u(\omega)$ is the low-pass filter with the upper cutoff frequency, ω_u, and $lp_l(\omega)$ is the low-pass filter with the lower cutoff frequency, ω_l.

It is standard terminology in the filtering literature to define an *ideal filter* as filters which cuts off sharply at some cutoff frequency. The power transfer function of an *ideal* low-pass filter satisfies

$$PTF^*_{lp}(\omega) \equiv \begin{cases} 1 & \text{if} \quad |\omega| < \omega_l \\ 0 & \text{if} \quad |\omega| \geq \omega_l \end{cases}, \tag{12}$$

the power transfer function of an ideal high-pass filter satisfies

$$PTF^*_{hp}(\omega) \equiv \begin{cases} 0 & \text{if} \quad |\omega| < \omega_l \\ 1 & \text{if} \quad |\omega| \geq \omega_l \end{cases}, \tag{13}$$

and the power transfer function of an ideal band-pass filter satisfies

$$PTF^*_{bp}(\omega) \equiv \begin{cases} 0 & \text{if} \quad |\omega| < \omega_l \\ 1 & \text{if} \quad \omega_l \leq |\omega| \leq \omega_u \\ 0 & \text{if} \quad |\omega| > \omega_u \end{cases}, \tag{14}$$

Figure 3.

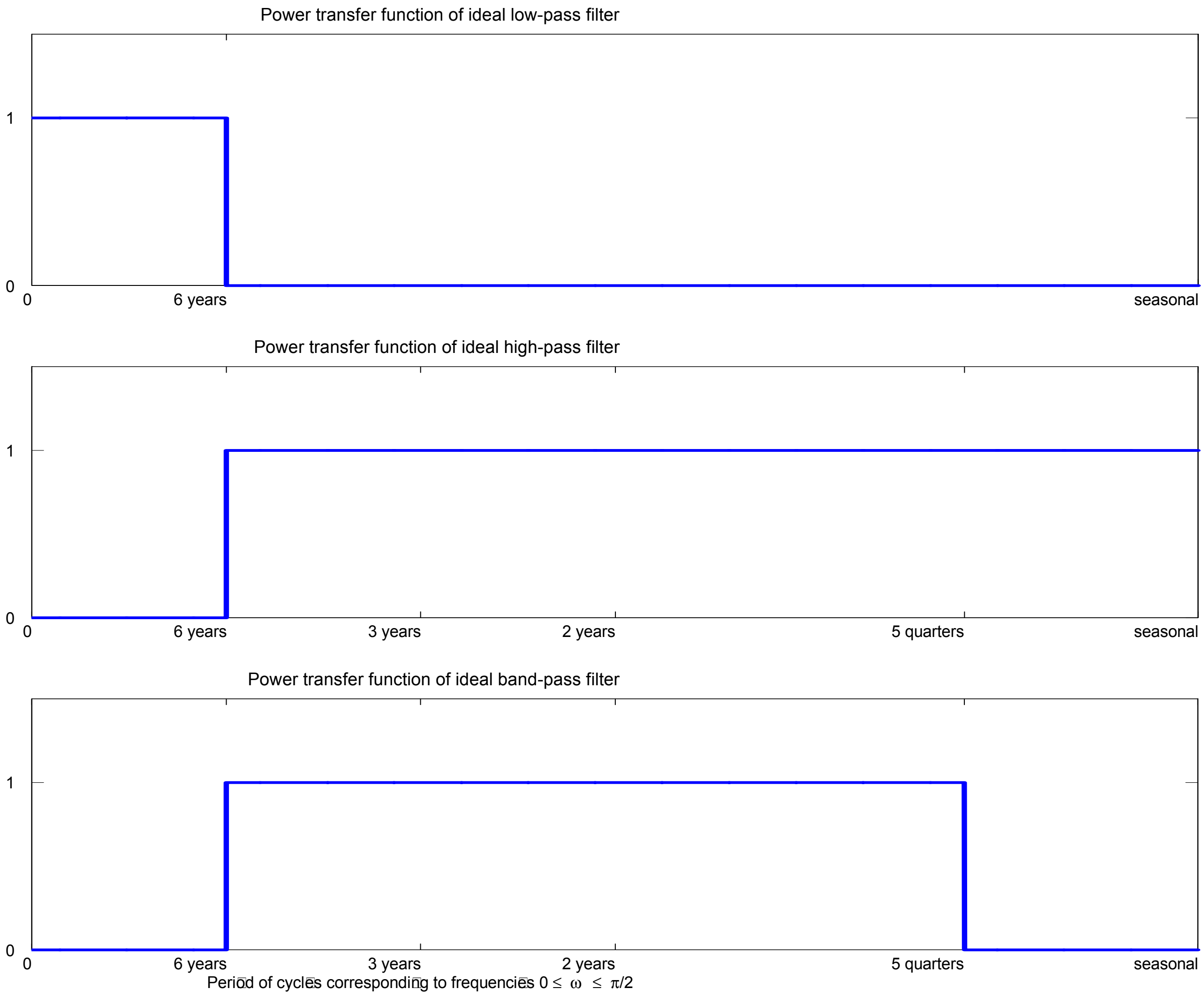

Figure 3: The power transfer function of (1) ideal low-pass filter, (2) ideal high-pass filter, and (3) ideal band-pass filter with cutoff frequencies corresponding to 6 years and 1.25 years

Many filters are approximations either to an ideal high-pass filter or to an ideal band-pass filter, thus there are two ways of defining the business cycle component:

1. The business cycle component can be defined by filtering the trend, τ_t, from the economic time series and interpreting the remaining components as the business cycle component. The widely used Hodrick-Prescott filter is often interpreted as an approximation to an ideal high-pass filter. High-pass filters are often used when filtering seasonally adjusted time series.

2. The business cycle component can also be defined by filtering the trend and the seasonal and irregular components from the economic time series, leaving the business cycle component c_t.

The power transfer function of an ideal low-pass filter is unity at zero frequency so that it allows all the power at zero frequency to pass the filter. That property is satisfied if the sum of the filter weights in the time domain sum to one

$$h(1) = \sum_{j=-\infty}^{\infty} h_j = 1, \tag{15}$$

and the frequency response function of the filter evaluated at frequency 0, is

$$h\,(0) = \sum_{j=-\infty}^{\infty} h_j e^{-i0j}, \tag{16}$$

which is equal to one if and only if the sum of the filter weights is one.

For an ideal high-pass filter, the jth weight of the high-pass filter is simply one minus the value of the jth weight of the low-pass filter.

If the sum of the weights in a symmetric low-pass filter is one, then the sum of the weights in the corresponding high-pass filter is zero and the power transfer function of an ideal high-pass filter is zero at zero frequency and it removes any unit root in an $I(1)$ time series.

Multiplying the spectrum in Figure 1 by an ideal band-pass filter gives the spectrum of the output process, Figure 4.

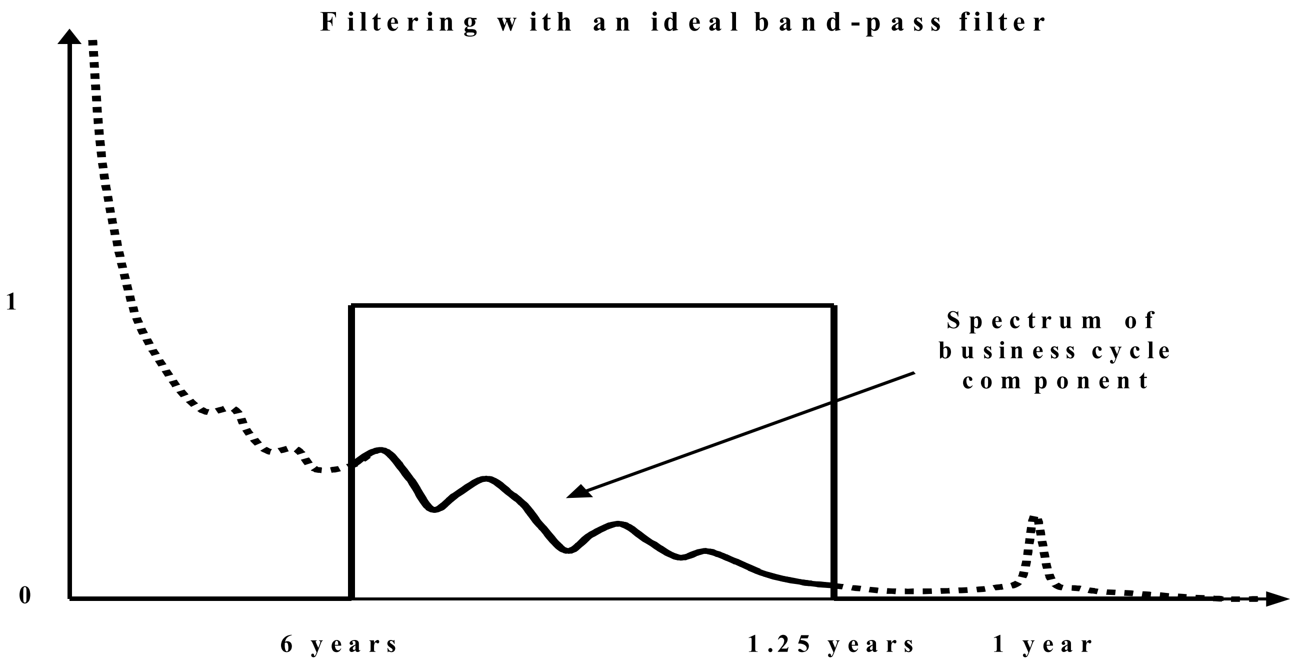

Figure 4: The effect of filtering with an ideal band-pass filter

Applying an ideal filter to the spectrum of any stochastic process which is not white noise has the consequence that the variance and autocovariance of the filtered process generally change relative to the input process

$$\int_0^\pi 2S_y(\omega)\, e^{i\omega\tau} d\omega = \gamma_\tau \neq \int_0^\pi PTF_{hp}^*(\omega) 2S_y(\omega)\, e^{i\omega\tau} d\omega. \tag{17}$$

The same is true for the cross-correlation between different time series.

The point is that a change in the measured population second moments of filtered and unfiltered time series cannot be interpreted as evidence of a distortionary effects of a filter, as do King and Rebelo (1993), Guay and St-Amant (1997), and Ehlgen (1998), since population moments change when filtering the time series, even when filtering with an ideal filter.[4]

2.5 Optimal Filtering

An optimal filter is one which is best in a certain sense and optimal filtering is about minimizing the mean square error of the business cycle component. Let $\{c_t^*\}_{t=1}^T$ be the "true" business cycle component of a time series $\{y_t\}_{t=1}^T$ and let $\{\hat{c}_t\}_{t=1}^T$ be its estimate. An optimal business cycle filter in the time domain, according to Wiener (1949), Whittle (1983), and King and Rebelo (1993), is one which minimizes the mean square error

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(\hat{c}_t - c_t^*)^2. \tag{18}$$

Let the true business cycle component be the result of filtering $\{y_t\}_{t=1}^T$ with a true business cycle filter (say, an ideal high-pass filter or an ideal band-pass filter), $c_t^* = C^*(L)\, y_t$, where $C^*(L)$ is the true business cycle filter in the time domain and let $\{\hat{c}_t\}_{t=1}^T$ be the result of filtering $\{y_t\}_{t=1}^T$ with some distorting business cycle filter, $\hat{c}_t = \hat{C}(L)\, y_t$, where $\hat{C}(L)$ is some estimated business cycle filter, then the metric (18) can be rephrased as

$$MSE = \frac{1}{T}\sum_{t=1}^{T}\left[\hat{C}(L)\, y_t - C^*(L)\, y_t\right]^2. \tag{19}$$

Optimal filtering is about choosing the filter, $\hat{C}(L)$, which minimizes the mean square error (19) for the time series y_t. This is a filtering or signal extraction problem.[5] Assume that y_t can be decomposed into a trend and a cyclical component,

$$y_t = \tau_t + c_t,$$

[4]The only exception to this rule is when filtering a white noise process since the variance of a white noise process is constant at all bands of frequencies.

[5]The signal extraction problem for the case where the two unobserved components are independent and stationary was solved independently by Wiener (1949) and by Kolmogorov in two papers published in French and German in 1939 and 1941, Bell (1984).

then Whittle (1983) shows that the optimal symmetric linear signal extraction filter of the cyclical component in the time domain satisfies

$$C(L) = \frac{g_c(L)}{g_\tau(L) + g_c(L)},$$

where $g_c(L)$ is the time domain autocovariance generating function of the true cyclical component and $g_\tau(L)$ is the time domain autocovariance generating function of the true trend component.

Since we do not know the autocovariance generating function of the true trend and business cycle component, we need to impose some identifying restrictions in order to estimate the cyclical component. It has been common to solve such signal extraction problems with the Kalman filter, Kalman (1960), or the Wiener-Kolmogorov filter, by formulating general unobserved component models and estimate the trend component and the cyclical component by imposing sufficient ad hoc identifying restrictions in order to estimate the unobserved components. King and Rebelo (1993) derive the optimal filter when both the cyclical component and the trend component are assumed to be generated by stationary ARMA-processes with known variances of the innovations which are assumed to be white noise and orthogonal for the two unobserved stochastic processes. Bell (1984) shows how to extend the signal extraction problem to non-stationary time series analyzed by Harvey (1985), Watson (1986), and Harvey and Jaeger (1993) and others. Kaiser and Maravall (2001) show how this can be done for a general unobserved component model, like (1), when the observed time series, y_t, follows a general ARIMA model.

Any such model based approach relies on arbitrary identifying assumptions about which we have no or very little a priori information. This gives rise to a major weakness with this approach, namely that it is difficult if not impossible to evaluate an unobserved component estimation procedure and it is difficult to compare different models.

2.6 Finite Time Approximations to Ideal Filters

There is another approach to optimal filtering developed by Koopmans (1974) based on a approximation to an ideal filter using a Fourier series and then using the inverse discrete Fourier transform to generate the filter weights in the time domain. This method is no less ad hoc than the model based approach since it is based on the ad hoc assumption, that the true business cycle filter is an ideal high-pass filter or ideal band-pass filter.

Koopmans (1974) showed how it is possible to construct an ideal low-pass filter in the time domain as an infinite dimensional symmetric linear time-invariant (LTI) filter

$$\tau_t^* = \sum_{j=-\infty}^{\infty} h_j^* y_{t-j}, \tag{20}$$

with the filter weights, h_j^*, satisfying $\sum_{j=-\infty}^{\infty} \left|h_j^*\right| < \infty$. The filter weights are found by the inverse discrete Fourier transform of the frequency response function of an ideal low-pass filter. Let the frequency response function of an ideal low-pass filter be

$$h^*(e^{-i\omega j}) = \sum_{j=-\infty}^{\infty} h_j^* e^{-ij\omega}, \tag{21}$$

with coefficients determined from the inverse discrete Fourier transform of the frequency response function

$$h_j^* = \frac{1}{2\pi}\sum_{-\pi}^{\pi} h^* e^{-i\omega j} e^{i\omega j} d\omega, \; j = 0, \pm 1, 2, ... \tag{22}$$
$$= \begin{cases} \omega_l/\pi & \text{for} \quad j = 0 \\ \sin(j\omega_l)/\pi j & \text{for} \quad j = \pm 1, 2, ... \end{cases}.$$

It is not possible to construct an ideal filter with only a finite number of observations but it is possible to construct a finite sample "optimal approximation" to the ideal low-pass filter as a finite dimensional moving average filter where the trend is given by the symmetric linear filter

$$\tau_t = \sum_{j=-K}^{K} a_j y_{t-j}. \tag{23}$$

There is a remarkable result saying that the optimal approximation to the ideal low-pass filter for given lag length K is constructed by simply truncating the ideal filter weights h_j^* from (22) at lag K[6] so that

$$a_j = \begin{cases} h_j^* & \text{for} \quad j = 0, \pm 1, 2, ..., K \\ 0 & \text{for} \quad j > K \end{cases}. \tag{24}$$

The power transfer function of a finite sample approximation to an ideal high-pass filter is plotted in Figure 5. The approximation to an ideal high-pass filter should improve as K increases. That is not necessarily the case for small changes in K, Pedersen (1999). The approximation does not mirror the ideal high-pass filter very well near the cutoff frequency for low values of K. The filter behaves like damped sinusoids for frequencies higher than ω_l. The poor approximation near the cutoff frequency is typical of Fourier series approximations to functions with discontinuities, known as the *Gibbs phenomenon*, Brockwell and Davis (1996).

3 The Distortionary Effect of Filters

A filter is distorting compared with an ideal filter when it cannot cut off sharply at any prespecified cutoff frequency. The distortionary effect of filters may take several forms. In general, a filter is distorting by passing (leaking) frequencies which it was supposed to attenuate and compressing frequencies which should pass the filter, Figure 6, therefrom the terms *leakage* and *compression*.[7]

[6] Koopmans (1974, ch. 6) and Baxter and King (1999).

[7] The term leakage is used in the spectral analysis literature. The term compression is - I believe - introduced by Baxter and King (1999).

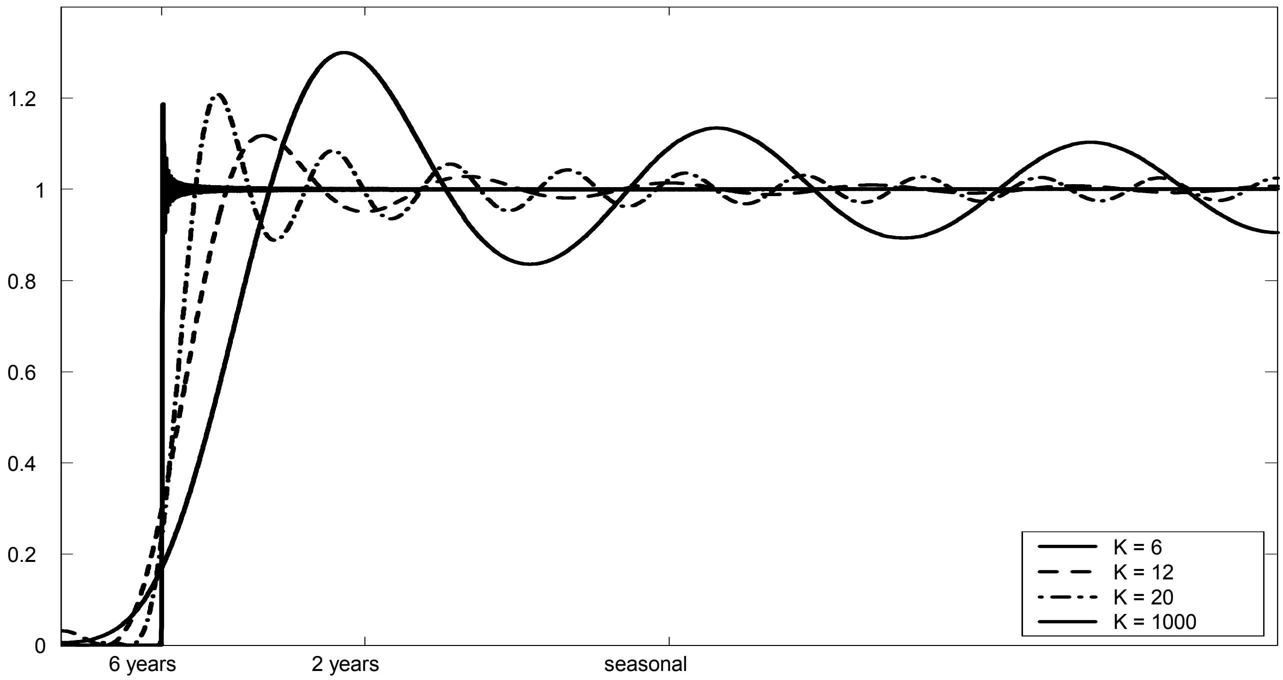

Figure 5: The power transfer function of a finite sample optimal approximation to an ideal high-pass filter with different truncation lags

The terms leakage and compression can be given a precise definition and measured in the frequency domain. The size of the leakage and compression of a filter is twice the absolute value of the area between the power transfer function of an ideal filter and the power transfer function of a distorting filter, which is approximately

$$\text{Leakage and compression} = \sum_{\omega \in W} \left| PTF\left(\omega\right) - PTF^{*}\left(\omega\right) \right| \cdot 2 \cdot \Delta\omega, \tag{25}$$

where $PTF^{*}(\omega)$ is the power transfer function of an ideal business cycle filter at frequency ω, where ω is n discrete equally spaced values between zero and π, $\omega \in W = (\omega_1 < \omega_2 < ... < \omega_n)$, with $\omega_1 = 0$ and $\omega_n = \pi$, $PTF(\omega)$ is the power transfer function of some distorting business cycle filter, and $\Delta\omega = \omega_i - \omega_{i-1}$, $i = 2, ..., n$.[8]

A special case of the distortionary effect of filters occurs if the power transfer function of a filter has cycles in which case the filter imposes spurious cycles on the filtered time series even when filtering a white noise process, the *Slutzky-effect*.

The Hodrick-Prescott filter (HP-filter) has been criticized by Harvey and Jaeger (1993), Cogley and Nason (1995), Park (1996), Guay and St-Amant (1997), Ehlgen (1998), and Kaiser and Maravall (2001) for inducing spurious cycles in filtered time series with the "typical spectral shape" of Granger (1966) and Cogley and Nason (1995) argue that the resulting "stylized facts" are rather "stylized artifacts" since they are merely the result of the filter and not of the time series itself.

[8]The first and last observation in the series $\Delta\omega$ is multiplied by 0.5.

By making a clear distinction between the effect of filtering with an ideal filter and with a distortionary filter, it is shown in Pedersen (2001) that the spuriousness critique rests on an inadequate definition of the Slutzky effect - a definition which has the unfortunate consequence that even an ideal high-pass filter induces a Slutzky effect.

There is no Slutzky effect in the HP filter when defining the Slutzky effect as cycles in the power transfer function of a filter. Besides, the Slutzky-effect is just a special case of a more general pattern of distortions and there is no reason to be more concerned about the Slutzky-effect than of any other pattern of distortions if the purpose of filtering is to compute business cycle stylized facts. A filter may be highly distorting without inducing spurious cycles for example by involving a dramatic re-weighting of frequencies compared with the ideal filter or it may be very little distorting even though there is a clear Slutzky-effect if the Slutzky-effect occurs at frequencies where the input time series has very little power. The leakage and compression of a filter is not a measure of the distortionary effect of a filter. The distortionary effect of a filter depends both on the power transfer function of the filter and on the spectrum of the filtered time series, that is, the leakage and compression at frequency ω is weighted by the relative power of the input series at frequency ω.

3.1 A Metric for Measuring the Distortionary Effect of Filters

It is not possible to measure the distortionary effect of a filter in the time domain when we only have a finite number of observations. This problem mirrors the impossibility of constructing an optimal filter in the time domain with a finite number of observations. With an infinite number of observations, the ideal cyclical component (defined as cycles with a period shorter than some prespecified cutoff frequency) is

$$C^{*}(L)\, y_t = \sum_{j=-\infty}^{\infty} \left(1-h_j^{*}\right) y_{t-j}, \tag{26}$$

with filter weights given by (22) and the estimated or finite sample approximation to the ideal cyclical component is

$$\hat{C}(L)\, y_t = \sum_{j=-K}^{K} \left(1-h_j^{*}\right) y_{t-j}. \tag{27}$$

The mean square error

$$\begin{aligned} MSE &= \frac{1}{T}\sum_{t=1}^{T}\left(\hat{C}(L)\, y_t - C^{*}(L)\, y_t\right)^2 \\ &= \frac{1}{T}\sum_{t=1}^{T}\left[\sum_{j=-K}^{K}\left(1-h_j^{*}\right) y_{t-j} - \sum_{j=-\infty}^{\infty}\left(1-h_j^{*}\right) y_{t-j}\right]^2 \end{aligned} \tag{28}$$

includes a sum of infinite terms.

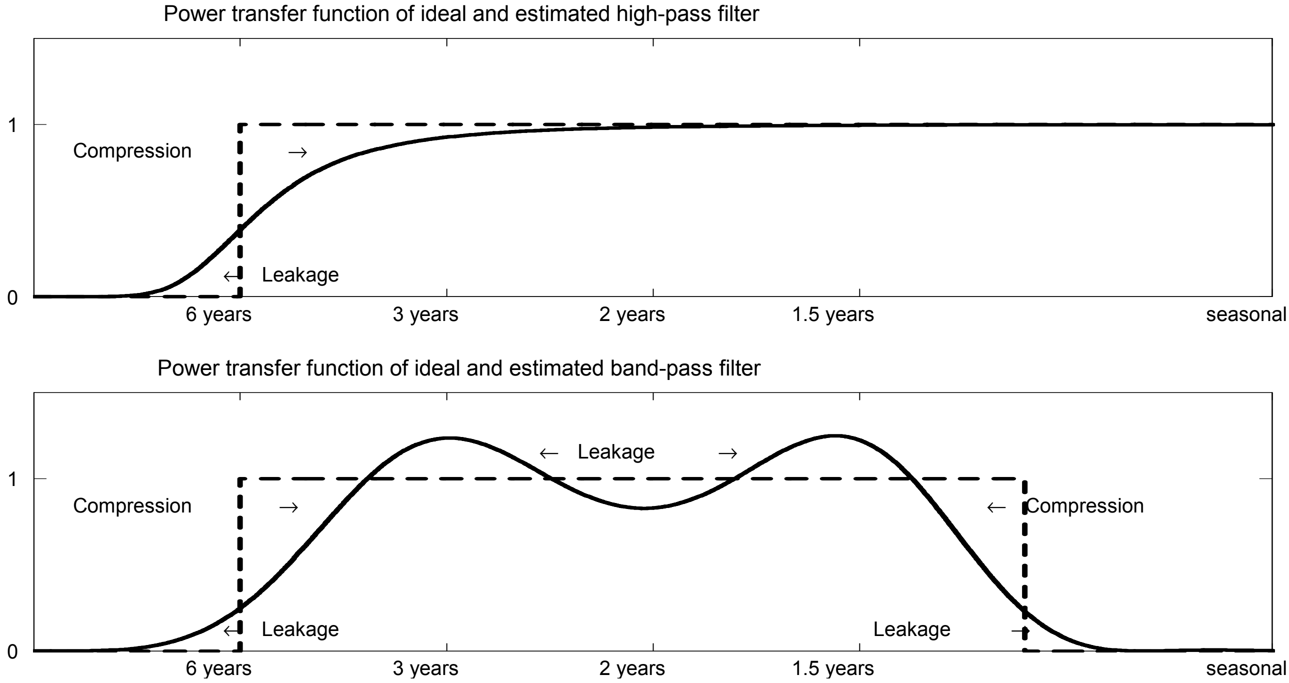

Figure 6: The power transfer function of two ideal business cycle filters and of two distorting filters

Let the ideal high-pass filter (13) or ideal band-pass filter (14) define the ideal business cycle filter. Let $PTF^*(\omega)$ be the power transfer function of such an ideal business cycle filter at frequency ω, where $\omega \in W = (\omega_1 < \omega_2 < ... < \omega_n)$, with $\omega_1 = 0$ and $\omega_n = \pi$, and let $PTF(\omega)$ be the power transfer function of some distorting business cycle filter. The spectrum of the true business cycle component is given by

$$S_c^*(\omega) = PTF^*(\omega) \cdot 2 \cdot S_y(\omega)$$

and the spectrum of the estimated business cycle component is

$$S_c(\omega) = PTF(\omega) \cdot 2 \cdot S_y(\omega).$$

The area under the integral $\int_0^\pi 2S_c^*(\omega)\,d\omega = \int_0^\pi PTF^*(\omega) \cdot 2 \cdot S_y(\omega)\,d\omega$ is the variance of the true business cycle component.

By using a discrete approximation of frequencies, the distortionary effect of a filter is measured as the sum[9] of the absolute value of the difference between the spectrum of the true business cycle component and the spectrum of the distorted business cycle component at frequency ω multiplied by the size of the grid on ω: $\Delta\omega = \omega_i - \omega_{i-1}$, $i = 2, ..., n$[10]

$$Q = \sum_{\omega \in W} |S_c(\omega) - S_c^*(\omega)| \cdot \Delta\omega, \qquad (29)$$

[9] Actually an integral but we approximate the metric using a discrete set of values of ω.
[10] The first and last observation in the series $\Delta\omega$ is multiplied by 0.5.

or rephrased as

$$Q = \sum_{\omega \in W} |PTF(\omega) - PTF^{*}(\omega)| \cdot 2 \cdot S_y(\omega) \cdot \Delta\omega, \tag{30}$$

It is seen from (30) that the distortionary effect of a filter depends on the power spectral density of the input process just as the MSE depends on y_t in (19). This is an important insight since it implies that we cannot in general measure the distortionary effect of a filter by the power transfer function and minimizing the leakage and compression of a filter is not in general equivalent to minimizing the distortionary effect of a filter.

If the spectrum is close to zero at frequencies where a filter is highly distorting, then the distortionary effect of the filter is small and vice versa. Most macroeconomic time series which have the "typical spectral shape" of Granger (1966) have most of their power at the very low growth frequencies so that the deviation of the power transfer function of a distorting filter from that of an ideal business cycle filter at low frequencies results in greater distortions than a quantitatively similar deviation at high frequencies.

It implies that different filters may be optimal for different time series or for the same time series for different countries. Specifically, the smoothing parameter of the Hodrick-Prescott filter, λ, should be determined as the value which minimizes the distortionary effect of the HP-filter for each country, just as the cut-off frequency of Baxter and King's (1999) band pass filter could depend on the country.

The distortionary effect of a filter is twice the absolute value of the area between the ideal filtered spectrum and the distortionary filtered spectrum, as illustrated for the case of the HP-filter in panel 2 in Figure 7.

The Q-statistics, (30), measures approximately twice the absolute value of the area between the two filtered spectra, panel 3 in Figure 7. Twice the absolute value of the area has the interpretation of the variance of the error as the MSE in (18).

The weights at frequency ω in (30) can be normalized with the effect that they sum to unity

$$Q_w = \sum_{\omega \in W} |PTF(\omega) - PTF^{*}(\omega)| \cdot v(\omega), \tag{31}$$

where the weights, $v(\omega)$, are determined as the ratio of the power spectral density of the input process, $S_y(\omega)$, at frequency ω to the variance of the series. The discrete sum of twice the power spectral density function over the frequencies $\omega \in W$ is approximately the variance

$$\sum_{\omega \in W} 2 \cdot S_y(\omega) \cdot \Delta\omega \simeq \gamma_0,$$

and the filter weights at frequency ω are constructed as follows:

$$v(\omega) = \frac{2 \cdot S_y(\omega) \cdot \Delta\omega}{\sum_{\omega \in W} 2 \cdot S_y(\omega) \cdot \Delta\omega}.$$

A new set of filter weights are constructed for each stochastic process. If the filter weights are equal, then this corresponds to filtering a white noise process where all frequencies have equal weight.

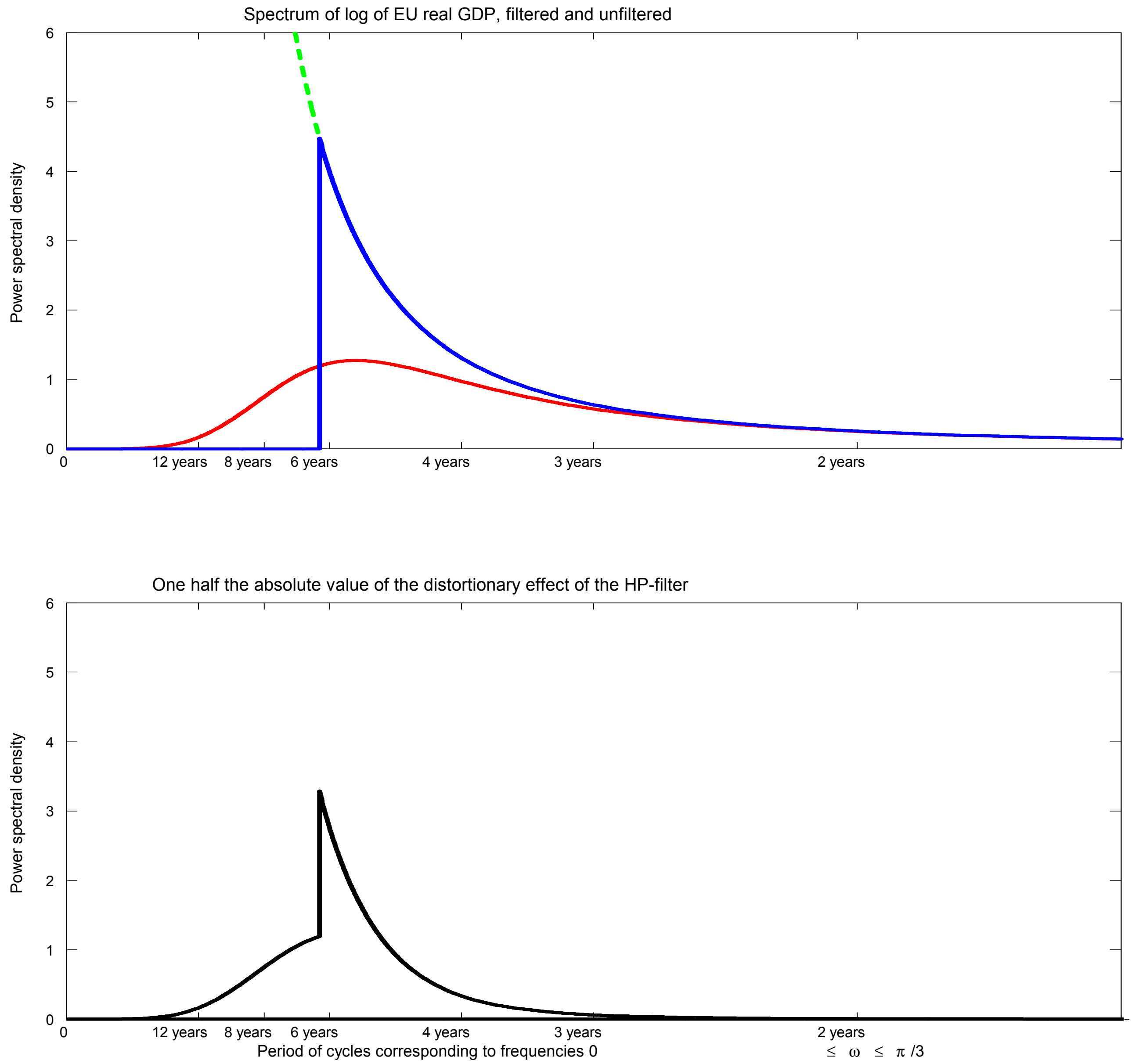

Figure 7: The distortionary effect of the HP-filter

4 Defining and Measuring the Business Cycle Component

The first and most important problem in empirical business cycle research is that of defining the business cycle component. There seems to have emerged a consensus about a *statistical definition* of the business cycle component in the macroeconomic and general equilibrium business cycle literature where a vast majority of researchers define the business cycle component either as an ideal high-pass filter or as an ideal band-pass filter. The two most commonly used filters are the Hodrick-Prescott filter (HP-filter) and Baxter and King's (1999) band-pass filter. The HP-filter has been interpreted as an approximation to an ideal high-pass filter by Prescott (1986) and King and Rebelo (1993) and Baxter and King's (1999) band-pass filter is constructed as an approximation to an ideal band-pass filter. Those filters are also widely used in empirical business cycle research in many international organizations, central banks, and government agencies.

Furthermore, a consensus seems to have developed about the duration of the business cycle component. A majority of researchers have defined the business cycle component as growth cycles with a duration less than eight years or between 1.5 and eight years. The Hodrick-Prescott (1980, 1997) filter with a value of the smoothing parameter $\lambda = 1600$ with quarterly data has been interpreted as an approximation to an ideal high-pass filter eliminating frequencies of 32 quarters or greater, Prescott (1986) and King and Rebelo (1993) and Baxter and King (1999) extract cycles with a duration between 1.5 and eight years.

Researchers relying on other methods seem to share this view of the duration of growth cycles. Englund, Persson, and Svensson (1992) and Hassler, Lundvik, Persson, and Söderlind (1994) use band-pass filters in the frequency domain to extract cycles with a duration between three and eight years, Canova (1998) defines the business cycle component as cycles with periods between 1.25 and 7.5 years, Burnside (1998) defines the business cycle component as cycles with periods between 1.5 and eight years, and King and Rebelo (1999) define the business cycle component as "those fluctuations in economic time series that have a periodicity of eight years or less."

However, there is no independent measurements underlying that view. The myth of an eight year business cycle component is often based on Burns-Mitchell (1946) measurements of the duration of *classical business cycles* which is measured in the *level* of aggregate economic activity and which is influenced by economic growth. A long period of economic growth is defined by Burns and Mitchell (1946) as an *intracycle* trend and measured as a part of the expansion phase. By contrast, the duration of *growth cycles* should be measured after detrending the data, Pedersen (2002).

It has not been uncommon to confuse the two concepts of "business cycles", and Canova (1999) compares the selection of turning points of growth cycles derived with a variety of filters with turning points in classical business cycles. That is of course an invalid testing procedure because statistics of two very different concepts of the "business cycle" are compared.

When the distortionary effect of a filter depends both on the transfer function of the filter and on the time series being filtered, then we need measurements on the duration of growth cycles in each country in order to determine the cutoff frequency of an ideal business cycle filter for each country. The duration of growth cycles can be measured as the time distance from peak to peak or from trough to trough in the business cycle component after detrending the data.

Turning points are identified in classical business cycles by Burns and Mitchell (1946) and their methods is continued by the NBER and since 1980 by its business cycle dating committee. The major difficulty with replicating the NBER method and use it for different countries is that the method is not well defined and relies on judgment by the researcher. Bry and Boschan (1971) developed a computer algorithm for the determination of turning points in individual time series which can be used to determine turning points in business cycles based on quarterly real GDP. The procedure has been modified based on modern filtering theory and programmed in

MATLAB, Mathworks (2002).[11] The computer algorithm is approximating the NBER selection of turning points based on a single reference series, quarterly real GDP. The selected turning points for the U.S. are close to the turning points selected by NBER, Elmer and Pedersen (2002). When identifying turning points in detrended time series, I use the same definition and the same programmed method for determining turning points as when analyzing classical business cycles, as do Mintz (1969, 1972, 1974) and Pagan (1997).

The computer algorithm is approximating the NBER selection of turning points based on a single reference series, real GDP. First, tentative peaks (P) and troughs (T) are determined in a highly smoothed time series, and subsequently the dating of turning points is refined based on less smoothed series. Each phase (peak to trough or trough to peak) must have a duration of at least six months and a cycle from peak to peak or from trough to trough must have a duration of at least 15 months. By using this computer algorithm, we avoid the difficulties of replicating the analysis of the NBER Business Cycle Dating Committee. Secondly, the same criteria for selecting turning points are applied for all countries. Finally, we avoid judgement, so that scientific replication is made possible.

The duration of growth cycles cannot be measured directly since the business cycle component is unobservable and it is necessary to detrend the time series of quarterly real GDP be3fore measuring the duration of growth cycles but by detrending the data, we define the business cycle component which we want to measure. That seems like an unsolvable problem. It is not, however, if economic fluctuations are *not* equally spaced over frequencies, that is, if the peaks and troughs of growth cycles are concentrated within a specific band of frequencies. That should not be a controversial assumption. It is widely believed that even though business cycles are far from periodic, quarterly real GDP does have a concentration of power at specific business cycle frequencies, Granger (1966), Sargent (1987), and King and Watson (1996). If that is the case, then it might be possible to remove low frequency fluctuations with a high-pass filter and measure the duration of growth cycles.

The initial detrending is done with the HP-filter with a low cutoff frequency. The cutoff of the HP-filter is not sharp, it passes frequencies which it was supposed to suppress, and it removes part of fluctuations which was supposed to pass the filter. The measurement of the duration of the true business cycle component is thus distorted but the bias goes both ways. The leakage biases the measurement towards longer growth cycles and the compression biases the measurement towards shorter growth cycles. It is not known which effect will dominate. However, in order to minimize the risk of removing too much of growth cycle frequencies and expose the results to the criticism of being biased towards shorter cycles by construction, I measure the duration of the business cycle component after filtering with cutoff frequencies corresponding to 20, 18, 16, 14, 12, and 10 years.

The critical assumption is that economic fluctuations are not equally spaced over frequencies. If the variance of a stationary economic time series is equally spaced over frequencies (like white noise), then it would be impossible to apply this measurement method. If such a process was filtered with an ideal high-pass filter with a cutoff frequency corresponding to, say, 14 years, then the cyclical component would display cycles with a duration up to 13.5 years, filtering with a cutoff frequency corresponding to 12 years would generate cycles with a duration up to 11.5 years etc. I do not find evidence of such a pattern. By contrast, the identification of

[11] This is joint work with Anne Marie Elmer, Elmer and Pedersen (2002).

turning points in growth cycles is in general quite robust to the definition of the business cycle component and the initial detrending, Table 3.

In order to detrend with the HP-filter with cutoff frequencies corresponding to 20, 18, 16, 14, 12, and 10 years, I need first to compute the optimal value of λ in the HP-filter for each of the countries and for each of the cutoff frequencies. The determination of λ^* follows the steps:

1. An $ARIMA(p, d, q)$-model is fitted to the logarithm of quarterly real GDP for each country. All combinations of models up to $d = 0, 1$ and $p = q = 4$ are estimated, and the model with the best fit is chosen. The ARMA-model is estimated with RATS, Estima (2000), using a Gauss-Newton algorithm. Let an $ARMA(p, q)$ process take the form

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q}, \tag{32}$$

   with its transfer function representation given by

$$y_t = \frac{1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q}{1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p} \cdot \varepsilon_t \tag{33}$$

2. The theoretical spectrum is computed with a version of the autoregressive spectral estimation method of Parzen (1969) and Berk (1974), which allow for general ARMA-processes, as in Hamilton (1994, ch. 6). The frequency response function is given by

$$h\left(e^{-i\omega}\right) = \frac{1 + \theta_1 e^{-i\omega} + \theta_2 e^{-2i\omega} + ... + \theta_q e^{-qi\omega}}{1 - \phi_1 e^{-1i\omega} - \phi_2 e^{-2i\omega} - ... - \phi_p e^{-pi\omega}}$$

   and the spectrum of y_t is given by

$$\begin{aligned} S_y(\omega) \quad = \quad & \frac{1}{2\pi} \left(\frac{1 + \theta_1 e^{-i\omega} + \theta_2 e^{-2i\omega} + ... + \theta_q e^{-qi\omega}}{1 - \phi_1 e^{-1i\omega} - \phi_2 e^{-2i\omega} - ... - \phi_p e^{-pi\omega}} \right) \\ & \cdot \left(\frac{1 + \theta_1 e^{i\omega} + \theta_2 e^{2i\omega} + ... + \theta_q e^{qi\omega}}{1 - \phi_1 e^{i\omega} - \phi_2 e^{2i\omega} - ... - \phi_p e^{pi\omega}} \right) \cdot \sigma^2 \end{aligned} \tag{34}$$

3. The optimal value of the smoothing parameter λ in the HP-filter is computed as the value, λ^*, which minimizes (31) for the cutoff frequencies corresponding to 20, 18, 16, 14, 12, and 10 years of an ideal high-pass filter.

The optimal values of λ are reported in Table 1 and 2. It is worth noting that the standard value, $\lambda = 1600$, comes close to being optimal when defining the cyclical component as growth cycles with a duration shorter than nine years and not eight years as is a standard interpretation. A value of $\lambda = 1000$ comes closer to being optimal for most countries, when the business cycle component is defined as cycles with a duration shorter than eight years. Spain, EU, and Finland have generally lower values corresponding to longer growth cycles. It is to be expected that the EU average has longer cycles than the individual countries since the economic fluctuations in individual EU countries have been far from perfectly correlated in the period 1960-2002.

Table 1: The minimum distorting value of λ in the HP-filter

	EURO	EU	BEL	DNK	FIN	FRA
3 years	18	18	32	22	12	24
4 years	59	52	83	67	52	71
5 years	145	115	185	160	125	165
6 years	310	230	360	330	250	340
7 years	580	420	650	610	480	620
8 years	1000	710	1090	1040	820	1050
9 years	1610	1140	1720	1660	1340	1680
10 years	2460	1750	2600	2530	2070	2550
11 years	3625	2575	3775	3700	3100	3725
12 years	5125	3700	5325	5225	4450	5275
13 years	7050	5150	7300	7200	6200	7250
14 years	9500	7000	9800	9700	8400	9800
15 years	12600	9400	12900	12700	11200	12800
16 years	16300	12400	16600	16500	14700	16700
18 years	20400	26200	26600	26400	24000	26800
20 years	31800	39800	40400	40200	37000	40800

Table 2: The minimum distorting value of λ in the HP-filter

	UK	ITA	NLD	SPA	SWE
3 years	18	18	25	11	19
4 years	56	58	73	34	58
5 years	135	145	170	90	140
6 years	290	310	340	200	300
7 years	540	580	630	390	560
8 years	940	1000	1060	690	960
9 years	1530	1600	1690	1160	1550
10 years	2360	2460	2560	1850	2390
11 years	3475	3625	3750	2800	3525
12 years	4975	5150	5300	4075	5000
13 years	6850	7100	7250	5750	6900
14 years	9300	9600	9800	7900	9400
15 years	12300	12600	12800	10600	12300
16 years	16000	16400	16600	14000	16100
18 years	25800	26400	26600	23200	25800
20 years	39400	40400	40600	36000	39400

4.1 The Duration of Growth Cycles

I have no knowledge of other studies asking the question about the duration of *growth cycles*. There is a large literature on the duration of *classical business cycles*, Burns and Mitchell (1946), Zarnowitz (1992), Watson (1994), Diebold and Rudebusch (1990, 1992, 1993, 1999).

The duration of growth cycles in each country is measured as the number of quarters from peak to peak or from trough to trough in quarterly real GDP after removing low frequency movements with the HP-filter.

There is an element of judgment in the determination of the maximum duration of growth cycles since there is no true way of detrending before identifying peaks and troughs of the growth cycle. Fortunately, the results are quite robust to the detrending method for many countries, Table 3. The maximum duration of any growth cycle in EU, Eurozone, Belgium, Finland, France, Italy, Sweden, and the UK is almost independent of the cutoff frequency of the HP-filter. We only find a large difference for Spain and with a difference of three quarters for Denmark and Netherlands.

For Denmark, we get the counterintuitive result that we identify a longer growth cycle with the higher cutoff frequencies than with the one corresponding to 16 years, and the longest cycle is chosen. The longest growth cycle in the Netherlands is 24 quarters when we filter with a cutoff frequency corresponding to 16, 14, and 10 years but it is 27 quarters when filtering with a cutoff frequency corresponding to 12 years. The value 24 quarters is chosen. The biggest difference between the maximum duration of growth cycles is found in Spain, where it is 46 quarters for the 20 year cutoff frequency but 26-27 quarters for the 16-18 year cutoff frequency and 19 quarters for the 14, 12, and 10 year cutoff frequency. The value 26 quarters is chosen.

I have in general chosen the growth cycle corresponding to the 14-16 year cutoff frequencies in Table 3, when in doubt, the only exception being the Netherlands. The minimum distorting value of λ in the HP-filter is around 16000 for a cutoff frequency corresponding to 16 years. The HP-filter with $\lambda = 16000$ leaves 92 percent of the 8-year cycle but it also includes 68 percent of 12-year cycles, which means that the measurements are biased heavily towards longer cycles. This business cycle component is in the following defined as growth cycles with a duration shorter than or equal to 6.25 years for the EU, 3.5 years for the Eurozone, 4 years for Belgium, 5.5 years for Denmark, 7 years for Finland, 4.25 years for France, 4.5 years for Italy, 6 years for the Netherlands, 6.5 years for Spain, 4.5 years for Sweden, and 5.25 years for the United Kingdom, Table 4.

Table 3: Maximum duration of growth cycles with cutoff frequency corresponding to 20, 18, 16, 14, 12, and 10 years, measured in quarters

	20 years	18 years	16 years	14 years	12 years	10 years
EU	26	26	25	25	25	25
Eurozone	13	14	14	13	13	13
Belgium	16	16	16	16	16	16
Denmark	24	24	19	22	22	22
Finland	28	28	29	28	28	28
France	17	17	17	16	16	16
Italy	18	18	18	18	18	18
Netherlands	27	27	24	24	27	24
Spain	46	27	26	19	19	19
Sweden	18	18	18	18	18	16
UK	21	21	21	21	21	21

Table 4: Definition of business cycle component and the optimal value of λ in the HP-filter

	Cutoff frequency Quarters	Optimal value of λ in HP-filter
EU	25	270
Eurozone	14	35
Belgium	16	85
Denmark	22	235
Finland	28	475
France	17	90
Italy	18	95
Netherlands	24	345
Spain	26	280
Sweden	18	90
UK	21	165

The large difference between the EU and the Eurozone countries may be caused by the different time periods covered. The EU data are for 1960.1-2002.2 and the Eurozone data are for the period 1991.1-2002.2.

The growth cycles are all shorter than the usual eight year definition of the business cycle component and similar results are found for the US, Canada, Japan, Australia, Austria, Korea, and Norway, Pedersen (2002).

A large number of empirical studies of international business cycles use the same value, $\lambda = 1600$, for all countries, Blackburn and Ravn (1991a, 1992), Backus and Kehoe (1992), Backus, Kehoe and Kydland (1992, 1993, 1995), Danthine and Donaldson (1993), Fiorito and Kollintzas (1994), OECD (2002) and others, but the distortionary effect of a filter depends both on the time series being filtered as well as on the power transfer function of the filter so different values of λ should be chosen for different countries - and in particular when growth cycles have different durations across countries. The optimal values of λ for the different countries are reported in Table 4.

5 Alternative Linear Business Cycle Filters

We have now the three ingredients for comparing alternative detrending techniques: (1) a metric for measuring the distortionary effect of filters, (2) an ad hoc statistical definition of the true business cycle component, and (3) an estimate of the business cycle component, the estimated spectrum of real GDP for each of the 11 countries and aggregates.

The filters studied are all linear, time-invariant filters: the first difference filter (FD), Blanchard and Fischer's (1989) $ARIMA(1,1,2)$-filter (BLFI), the Hodrick-Prescott filter (HP) with $\lambda = 1600$ and with the optimal value λ^* for each country, a finite sample optimal approximation to an ideal high-pass filter with the Baxter-King-constraint that the sum of the coefficients sum to zero (BKhp), Baxter and King's (1999) band-pass filter (BKbp), and a band-pass filter based on the Hodrick-Prescott filter (HPbp).[12]

5.1 First Difference Filter

An often used method for transforming $I(1)$-series to stationary time series is to take the first difference of the logarithm of the series. The first difference filter attenuates the low frequencies and amplifies higher frequencies. The power transfer function of the first difference filter is computed as follows

$$\begin{aligned} H_{\Delta}(\omega) &= \left(1 - e^{-i\omega}\right)\left(1 - e^{i\omega}\right) = 1 - e^{-i\omega} - e^{i\omega} + 1 \\ &= 2 - 2\cos(\omega). \end{aligned} \quad (35)$$

5.2 Blanchard and Fischer's ARIMA Filter

Blanchard and Fischer (1989) propose to estimate the trend component as an $ARIMA(1,1,2)$-model and they estimate the following model for U.S. quarterly log of real GDP for the period 1947:4-1987:2

$$\Delta y_t = .006 + .24\Delta y_{t-1} + \varepsilon_t + .08\varepsilon_{t-1} + 0.24\varepsilon_{t-2}.$$

Blanchard and Fischer's filter can be formulated as a sequential application of first, the first difference filter, and next the trend is an $ARMA(1,2)$-process. The power transfer function of the first difference filter is given by (35).

The frequency response function of the trend component, the $ARMA(1,2)$-filter

$$G(\omega) = \frac{(1 + .08e^{-i\omega} + .24e^{-2i\omega})}{(1 - .24e^{-i\omega})},$$

so the frequency response function of the business cycle component is

$$C(\omega) \equiv 1 - G(\omega) = 1 - \frac{(1 + .08e^{-i\omega} + .24e^{-2i\omega})}{(1 - .24e^{-i\omega})},$$

and the power transfer function of the business cycle component is

$$|C(\omega)|^2 = \left|1 - \frac{(1 + .08e^{-i\omega} + .24e^{-2i\omega})}{(1 - .24e^{-i\omega})}\right|^2,$$

The power transfer function of the business cycle component of the $ARIMA(1,1,2)$-filter is likewise the product of the power transfer function of the first difference filter and the power transfer function of the $ARMA(1,2)$-filter

$$H_{BLFI}(\omega) = \left|1 - \frac{(1 + .08e^{-i\omega} + .24e^{-2i\omega})}{(1 - .24e^{-i\omega})}\right|^2 \cdot [2 - 2\cos(\omega)], \quad (36)$$

[12] Details on the filters and on the computation of the power transfer function are given in Pedersen (1999), available from the author upon request together with MATLAB-files.

5.3 The Hodrick-Prescott Filter

The Hodrick-Prescott filter computes a stochastic trend $\{\tau_t\}_{t=1}^{T}$ by minimizing the sum of squared deviations of a time series from its trend $(y_t - \tau_t)^2$ subject to the constraint that the sum of the squared second differences is not too large

$$\min_{\{\tau_t\}_{\tau=1}^{T}} \sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 . \tag{37}$$

The first term is a measure of "goodness-of-fit" and the second term is a measure of the "degree-of-smoothness" which penalizes decelerations in the growth rate of the trend component. Variations in the smoothing parameter λ alters the trade-off between the goodness-of-fit and the degree-of-smoothness.

Differentiating (37) with respect to τ_t, the cyclical HP filter is

$$C(L) \equiv \frac{\lambda(1-L)^2 \cdot (1-L^{-1})^2}{\lambda(1-L)^2 \cdot (1-L^{-1})^2 + 1} = \frac{\lambda L^{-2}(1-L)^4}{\lambda L^{-2}(1-L)^4 + 1}, \tag{38}$$

with the power transfer function as the square of the absolute value of the frequency response function[13]

$$H_{HP}(\omega) = \left| \frac{\lambda(1-e^{-i\omega})^2 \cdot (1-e^{i\omega})^2}{\lambda(1-e^{-i\omega})^2 \cdot (1-e^{i\omega})^2 + 1} \right|^2 = \left| \frac{4\lambda [1-\cos(\omega)]^2}{4\lambda [1-\cos(\omega)]^2 + 1} \right|^2 . \tag{39}$$

5.4 Optimal approximation to an Ideal High-Pass Filter

An optimal approximation to an ideal low-pass filter was defined by the filter weights given by (22). Baxter and King (1999) impose the constraint that the filter weights of the low-pass filter sum to one in order for the high-pass filter to have zero gain at zero frequency. In order to remove a unit root, the approximated low-pass filter is imposed the constraint that the sum of the weights is unity in which case the corresponding high-pass filter has zero gain at zero frequency. Baxter and King therefore impose the side constraint that $\alpha_j(\omega) = 0$ which implies that

$$\alpha_j = h_j + \theta, \quad j = 0, \pm 1, 2, ..., K \tag{40}$$

where θ is a constant which depends on the lag length K

$$\theta = \frac{1 - \sum_{j=-K}^{K} h_j}{2K+1}.$$

[13]The HP filter was first analyzed in the frequency domain by Singleton (1988), King and Rebelo (1989,1993) and Blackburn and Ravn (1991b).

The power transfer function of a moving average filter with K lags is given by

$$H_{MA}(\omega) = \Big| 1 - \left(\frac{1}{2K+1}\right) \cdot \alpha_K e^{-i\omega K} + \alpha_{K-1} e^{-i\omega(K-1)} + ... + \alpha_2 e^{-i2\omega} + \alpha e^{-i\omega} + 1 + \alpha_1 e^{i\omega} + ... + \alpha_K e^{i\omega K} \Big|^2$$

and with the coefficients given by (40). I report the results for $K = 12$, 16, and 20.

5.5 Baxter and King's Band-Pass Filter

Baxter and King's (1999) band-pass filter is constructed as the difference between two optimal approximated low-pass filters with the constraint that the filter weights sum to zero. Baxter and King recommend to use the value $K = 12$ for quarterly data. A higher value of K, should give a better approximation to an ideal filter but limits the number of observations available to the researcher. I report the results for $K = 12$, 16, and 20. The upper cut off frequency corresponds to five quarters and the lower is determined by the duration of the growth cycle, given in Table 4.

5.6 Band-Pass Filter based on the Hodrick-Prescott Filter

I construct a band-pass filter based on the Hodrick-Prescott filter as a cascaded HP-filter. First, the series is filtered with the HP-high-pass filter with a low cutoff frequency, and next the "cyclical component" is filtered with a HP low-pass filter with a high cutoff frequency.[14]

The power transfer function of this band-pass filter is the product of the power transfer function of the high-pass and the low-pass filters. The lower and upper values of λ in the two low pass filters are determined simultaneously as the minimum distorting values for each country.

5.7 Results

The distortionary effects of the filters are reported in Table 5 and 6. The results depend on the time series being filtered and therefore depend on the country. There are no general results.

However, the first difference filter (FD) is highly distorting and it involves a dramatic reweighting of frequencies. It furthermore induces a phase shift. Blanchard and Fischer's (1989) filter is also highly distorting, it has a clear Slutzky effect, and it also induces a phase shift.

[14] A Hodrick-Prescott band-pass filter can also be constructed as a linear combination of two low-pass filters. This sequential filter is slightly less distorting.

Table 5: The distortionary effect of alternative business cycle filters

	Eurozone	EU	Belgium	Denmark	Finland	France
First difference filter	0.111	0.068	2.480	0.605	0.155	35.385
Blanchard-Fischer	0.105	0.063	2.490	0.563	0.150	22.807
HP, $\lambda = 1600$	0.250	0.054	1.699	0.363	0.090	16.790
HP, $\lambda = \lambda^*$	0.044	0.025	0.514	0.174	0.061	4.959
BKhp, K=12	0.030	0.033	0.469	0.195	0.076	4.501
BKhp, K=16	0.026	0.022	0.359	0.143	0.067	3.346
BKhp, K=20	0.022	0.019	0.324	0.120	0.048	3.172
BKbp, K=12	0.034	0.031	0.542	0.197	0.077	4.313
BKbp, K=16	0.030	0.022	0.519	0.135	0.064	3.927
BKbp, K=20	0.025	0.019	0.419	0.135	0.051	3.324
HP-bandpass	0.054	0.029	1.135	0.233	0.069	7.928

Table 6: The distortionary effect of alternative business cycle filters

	Italy	Netherlands	Spain	Sweden	United Kingdom
First difference filter	10.176	23.135	4.730	0.154	0.200
Blanchard-Fischer	10.349	18.640	4.299	0.141	0.172
HP, $\lambda = 1600$	15.056	7.660	4.738	0.192	0.162
HP, $\lambda = \lambda^*$	4.311	4.523	2.263	0.053	0.061
BKhp, K=12	3.666	5.318	3.040	0.046	0.066
BKhp, K=16	2.989	4.347	2.210	0.037	0.046
BKhp, K=20	2.699	3.364	1.624	0.034	0.043
BKbp, K=12	3.545	5.684	2.808	0.044	0.062
BKbp, K=16	3.124	4.078	2.100	0.040	0.044
BKbp, K=20	2.774	3.734	1.646	0.035	0.046
HP-bandpass	5.101	6.758	2.382	0.067	0.077

It is worth noticing how the HP-filter with the optimal values of λ in general are much less distorting than the HP-filter with the standard value $\lambda = 1600$. this shows that it is worth the effort to compute the optimal value of λ.

The Baxter-King approximated high-pass filter (BKhp) with the recommended value $K = 12$ is often about as good as the HP-filter with the optimal value and the performance of BKhp improves considerably when increasing K to 20.

Baxter and King's (1999) band-pass filter is generally superior to the band-pass filter based on the HP-filters.

When comparing the distortionary effect of filters, we have so far neglected the endpoint problems of the HP filter which arise because the HP filter is an infinite dimensional moving average filter in the time domain. Baxter and King (1999) recommend to cut off the first and last three years of observations when using the HP filter to minimize the endpoint problems. This makes it comparable with BKhp and BKbp for $K = 12$. The distortionary effect of the endpoint problems can be computed with the metric (31).

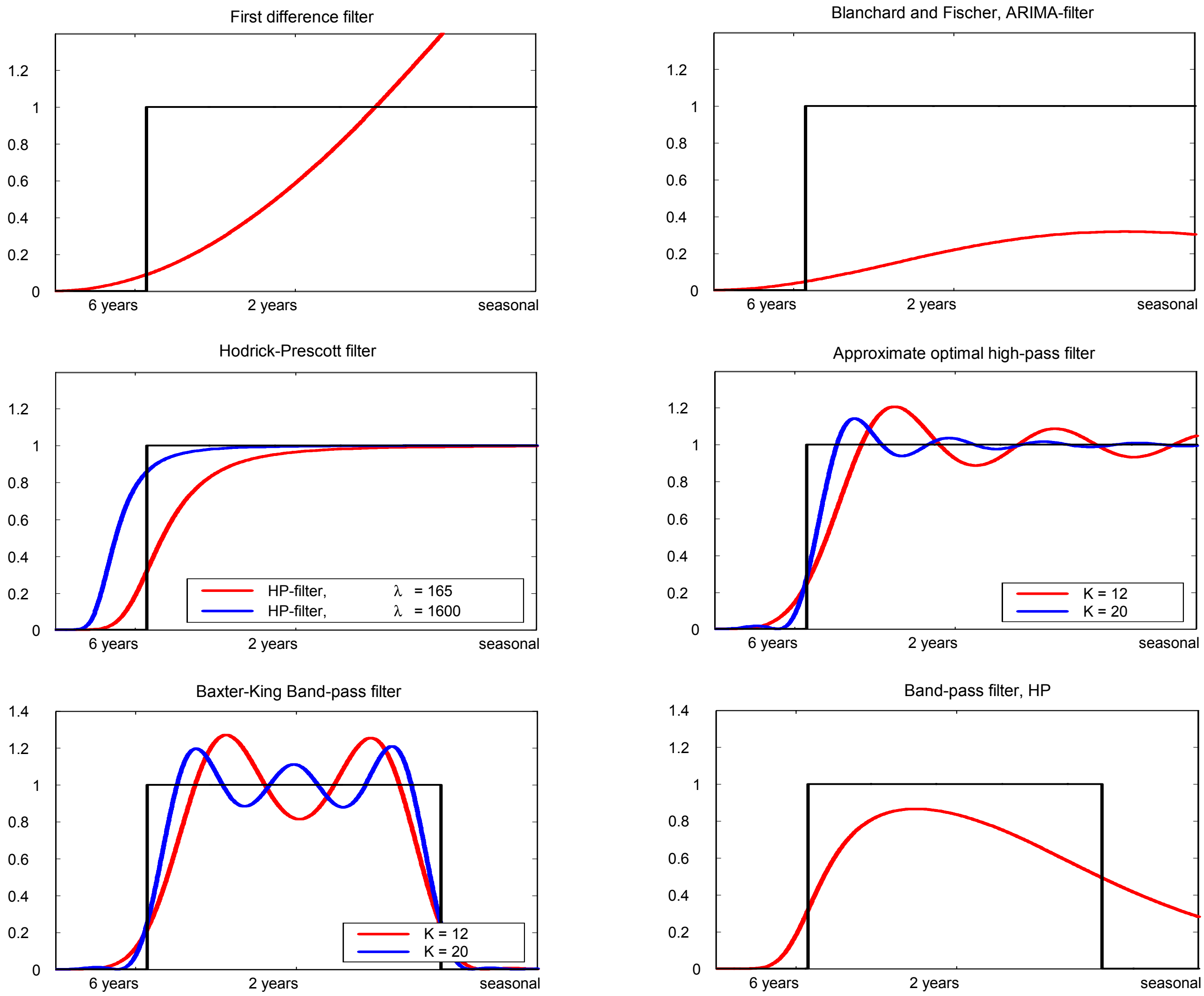

Figure 8: The power transfer function of filters for the UK

6 Summary and Conclusions

It has been common in modern business cycle research to detrend macroeconomic time series before analysis. It is possible to compare the distortionary effect of alternative filters for any given ad hoc definition of the "true" business cycle component. That should be a great help in choosing among the many available business cycle filters.

There are unfortunately no general results about the distortionary effect of filters apart from the fact that it depends on the time series being filtered and that different filters may be optimal for different countries or different time series. In particular, when filtering with the Hodrick-Prescott filter, the value of the smoothing parameter should be adjusted to the time series or country. It is not recommended to use the standard value $\lambda = 1600$ for all countries. The value $\lambda = 1600$ is close to being the optimal value for many countries when defining the business cycle component as cycles with a period up to nine years. Measurements of the duration of growth cycles in the EU countries show, however, that growth cycles are in general much shorter.

References

BACKUS, D. K., AND P. J. KEHOE (1992): "International Evidence on the Historical Properties of Business Cycles," *American Economic Review* 82, 864-88.

BACKUS, D. K., P. J. KEHOE, AND F. E. KYDLAND (1992): "International Real Business Cycles," *Journal of Political Economy* 100, 745-76.

BACKUS, D. K., P. J. KEHOE, AND F. E. KYDLAND (1993): "International Business Cycles: Theory vs. Evidence," *Federal Reserve Bank of Minneapolis. Quarterly Review* Fall 1993, 14-29.

BACKUS, D. K., P. J. KEHOE, AND F. E. KYDLAND (1995): "International Business Cycles: Theory and Evidence," in: Cooley (ed.): *Frontiers of Business Cycle Research.* Princeton. Princeton University Press.

BAXTER, M., AND R. G. KING (1999): "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series", *The Review of Economics and Statistics* 81, 575-593.

BELL, W. (1984): "Signal Extraction for Nonstationary Time Series," *The Annals of Statistics* 12, 646-64.

BERK, K. N. (1974): "Consistent Autoregressive Spectral Estimates",*The Annals of Statistics* 2, 489-502.

BLACKBURN, K., AND M. O. RAVN (1991a): "Contemporary Macroeconomic Fluctuations: An International Perspective," *University of Aarhus, Institute of Economics, Working paper no. 1991:12.*

BLACKBURN, K., AND M. O. RAVN (1991b): "Univariate Detrending of Macroeconomic Time Series," *University of Aarhus, Institute of Economics, Working paper no. 1991:22.*

BLACKBURN, K., AND M. O. RAVN (1992): "Business Cycles in the United Kingdom: Facts and Fictions," *Economica* 59, 383-401.

BLANCHARD, O. J., AND S. FISCHER (1989): *Lectures on Macroeconomics.* Cambridge, Ma. MIT Press.

BRY, G., AND C. BOSCHAN (1971): *Cyclical Analysis of Time Series. Selected Procedures and Computer Programs.* New York.

BURNS, A. F., AND W. C. MITCHELL. (1946): *Measuring Business Cycles.* New York.

BURNSIDE, A. C. (1998): "Detrending and Business Cycle Facts: A Comment," *Journal of Monetary Economics* 41, 513-32.

CANOVA, F. (1998): "Detrending and Business Cycle Facts," *Journal of Monetary Economics* 41, 475-512.

CANOVA, F. (1999): "Does Detrending Matter for the Determination of the Reference Cycle and the Selection of Turning Points?" *The Economic Journal* 109, 126-149.

Danthine, J. P., and J. B. Donaldson (1993): “Methodological and Empirical Issues in Real Business Cycle Theory,” *European Economic Review* 37, 1-35.

Diebold, F. X., and G. D. Rudebusch (1990): “A Non-Parametric Investigation of Duration Dependence in the American Business Cycle,” *Journal of Political Economy* 90, 596-616.

Diebold, F. X., and G. D. Rudebusch (1991): “Shorter Recessions and Longer Expansions,” Federal Reserve Bank of Philadelphia, *Business Review* June, 13-20.

Diebold, F. X., and G. D. Rudebusch (1992): “Have Postwar Economic Fluctuations Been Stabilized?” *American Economic Review* 82, 993-1005.

Diebold, F. X., and G. D. Rudebusch (1993): “Further Evidence on Business-Cycle Duration Dependence, in: Stock and Watson (eds.): *Business Cycles, Indicators, and Forecasting.* Cambridge, Cambridge University Press.

Diebold, F. X., and G. D. Rudebusch (1999): *Business Cycles. Durations, Dynamics, and Forecasting.* Princeton: Princeton University Press.

Elmer, A. M., and T. M. Pedersen (2002): “Identifying Turning Points in Business Cycles for 16 OECD Countries,” *Chapter 4 of Ph.D.-thesis*, Institute of Economics, University of Copenhagen.

Englund, P., T. Persson, and L. E. O. Svensson (1992): “Swedish Business Cycles: 1861-1988,” *Journal of Monetary Economics* 30, 343-71.

Engle, R. F., and C. W. J. Granger (1987): “Co-Integration and Error Correction: Representation, Estimation, and Testing,” *Econometrica*, vol. 55, no. 2, p. 251-76.

Fiorito, R., and T. Kollintzas (1994): “Stylized Facts of Business Cycles in the G7 from a Real Business Cycles Perspective,” *European Economic Review* 38, 235-69.

Frisch, R. (1933): “Propagation Problems and Impulse Problems in Dynamic Economics,” *Essays in Honor of Gustav Cassel.* London.

Granger, C. W. J. (1966): “The Typical Spectral Shape of an Economic Variable,” *Econometrica* 34, 150-61.

Guay, A., and P. St-Amant (1997): “Do the Hodrick-Prescott and Baxter-King Filters provide a Good Approximation of Business Cycles?” *Working Paper no. 53.* Université du Québec à Montréal.

Hamilton, J. D. (1994): *Time Series Analysis.* Princeton. Princeton University Press.

Harvey, A. C. (1985): “Trends and Cycles in Macroeconomic Time Series,” *Journal of Business and Economic Statistics* 3, 216-27.

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

eurostat

Hassler, J., P. Lundvik, T. Persson, and P. Söderlind (1994): "The Swedish Business Cycle: Stylized Facts over 130 Years," in: Bergström and Vredin (eds.): *Measuring and Interpreting Business Cycles.* Oxford.

Hodrick, R. J., and E. C. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking* 29, 1-16. Carnegie-Mellon working paper, 1980.

Kaiser, R., and A. Maravall (2001): *Measuring Business Cycles in Economic Time Series. Lecture Notes in Statistics 154.* Berlin: Springer-Verlag.

Kalman, R. E. (1960): "A New Approach to Linear Filtering and Prediction Problems," *Transactions ASME Journal of Basic Engineering* D 83, 35-45.

King, R. G.. and S. T. Rebelo (1989): "Low Frequency Filtering and Real Business Cycles," *Rochester Center for Economic Research. Working Paper no. 205.*

King, R. G., and S. T. Rebelo (1993): "Low Frequency Filtering and Real Business Cycles," *Journal of Economic Dynamics and Control* 17, 207-31.

King, R. G., and S. T. Rebelo (1999): "Resuscitating Real Business Cycles", in: Woodford and Taylor (eds.): *Handbook of Macroeconomics. Volume IA.* Amsterdam: North-Holland.

King, R. G., and M. W. Watson (1996): "Money, Prices, Interest Rates and the Business Cycle," *The Review of Economics and Statistics*, 35-53.

Koopmans, L. H. (1974): *The Spectral Analysis of Time Series.* New York. Academic Press.

Lucas, R. E. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4, 103-24. Reprinted in: Lucas: *Studies in Business Cycle Theory.* Cambridge, Ma. 1981.

Lucas, R. E. (1977): "Understanding Business Cycles," *Carnegie-Rochester Conference Series on Public Policy* 5.

Mathworks (2002): *MATLAB r13.*

Mintz, I. (1969): *Dating Postwar Business Cycles: Methods and their Application to Western Germany, 1950-67.* New York. NBER.

Mintz, I. (1972): "Dating American Growth Cycles," in: Zarnowitz (ed.): *The Business Cycle Today.* New York. NBER.

Mintz, I. (1974): "Dating United States Growth Cycles," *Explorations in Economic Research* 1, 1-113.

Moore, G. H., and V. Zarnowitz (1986): "The Development and Role of the National Bureau of Economic Research's Business Cycle Chronologies," in: Gordon (ed.): *The American Business Cycle. Continuity and Change.* Chicago.

Nelson, C. R., and C. I. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics* 10, 139-62.

OECD (2002): *OECD Economic Outlook*, April 2002. Paris.

Pagan, A. (1997): "Towards an Understanding of Some Business Cycle Characteristics," *The Australian Economic Review* 30, 1-15.

Park, G. (1996): "The Role of Detrending Methods in a Model of Real Business Cycles," *Journal of Macroeconomics* 18, 479-501.

Parzen, E. (1969): "Multiple Time Series Modelling," in: Krishnaiah (Ed.): *Multivariate Analysis, Vol. II.* New York. Academic Press.

Pedersen, T. M. (1999): *Spectral Analysis, Business Cycles, and Filtering of Economic Time Series. With MATLAB applications.* Manuscript. Institute of Economics, University of Copenhagen.

Pedersen, T. M. (2001): "The Hodrick-Prescott Filter, the Slutzky-Effect, and the Distortionary Effect of Filters," *Journal of Economic Dynamics and Control* 25, 1081-1101.

Pedersen, T. M. (2002): "How long are Growth Cycles?" *Chapter 5 of Ph.D. thesis. Institute of Economics, University of Copenhagen.*

Pedersen, T. M., and A. M. Elmer (2003): "International Evidence on the Connection between Business Cycles and Economic Growth," *Journal of Macroeconomics, forthcoming.*

Prescott, E. C. (1986): "Theory Ahead of Business-Cycle Measurement," *Carnegie-Rochester Conference Series on Public Policy* 25, 11-44.

Sargent, T. J. (1987): *Macroeconomic Theory.* 2nd. ed. New York: Academic Press.

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

Slutzky, E. (1937): "The Summation of Random Causes as the Source of Cyclic Processes," *Econometrica,* 105-46. In Russian, 1927.

Stock, J. H., and M. W. Watson (1999): "Business Cycle Fluctuations in U.S. Macroeconomic Time Series," in: Woodford and Taylor (eds.): *Handbook of Macroeconomics. Volume IA.* Amsterdam: North-Holland.

Söderlind, P. (1994): "Cyclical Properties of a Real Business Cycle Model," *Journal of Applied Econometrics* 9, S113-22.

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

Watson, M. W. (1994): "Business-Cycle Durations and Postwar Stabilization of the U.S. Economy," *American Economic Review* 84, 24-46.

Whittle, P. (1983): *Prediction and Regulation.* 2nd. Edition. Oxford. Basil Blackwell.

Wiener, N. (1949): *Extrapolation, Interpolation, and Smoothing of Stationary Time Series.* Cambridge, Ma. MIT Press.

Yule, G. U. (1927): "On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers," *Philosophical Transactions of the Royal Society of London A*, 226, 267-73.

Zarnowitz, V. (1992): *Business Cycles. Theory, History, Indicators, and Forecasting.* Chicago. University of Chicago Press for the NBER.

Data Appendix

The data are quarterly real GDP for the EU (1960:1-2002:2), the Eurozone (1991:1-2002:2) and for a number of individual EU-countries: Belgium (1980:1-2002:2), Denmark (1977:1-2002:2), Finland (1975:1-2002:2), France (1960:1-2002.2), Italy (1970:1-2002.3), the Netherlands (1960:1-2002.3), Spain (1970:1-2002:2), Sweden (1980:1-2002:2), and United Kingdom (1955:1-2002.3). The majority of the time series are from EcoWin. Some of the time series have been combined from different sources, presumably after changes in base year.

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Turning Points Detection"

A COMPARATIVE ASSESSMENT OF PARAMETRIC AND NON-PARAMETRIC TURNING POINTS DETECTION METHODS: THE CASE OF THE EURO-ZONE ECONOMY

By Jacques Anas and Laurent Ferrara[1]

Centre d'Observation Economique

[1]Centre d'Observation Economique, 27, Avenue de Friedland, 75382 Paris Cedex 08
Tel. : + 33 (0) 55 65 70 80 – Fax : + 33 (0) 55 65 70 74
E-mail: janas@ccip.fr, lferrara@ccip.fr

1 Introduction

In the last decade, the development of new modern tools, especially regarding non-linear modeling, has allowed a renewed interest in the empirical cyclical analysis and especially in de-trending techniques and cyclical turning point identification. The literature on this issue is gaining momentum. But there is a challenge ahead of us to disseminate those new tools among potential users and at the same time, to make them sufficiently transparent and understandable.

This paper is a point of view of practitioners, halfway between academic and official public research and private users in the business community[2]. Its aim is to present some point of views on the question of detecting and dating the global economic cycle and to offer some tentative comparisons on the main available methods used today. A comparative assessment on parametric and non-parametric methods for dating the euro-zone cycle will be presented.

2 Objectives

2.1 Distinction between dating and detecting

First of all, we want to insist on the separate issue of dating and detecting. To hazard a comparison, it is almost the same difference than between estimating and predicting.

Even if the past is by definition totally known, the timing of past turning points has to be estimated. This latter estimation can be carried out either by some reasonable non-parametric method (graphically or by using a pattern recognition algorithm), or by a model aiming at replicating this past reality. The "best" model has to be found to reach this objective, which means the best specification able to reproduce this precise stylized fact. But to test the validity of the model, we need to know this stylized fact, *i.e.* the past chronology of turning points. It seems therefore that there is an intrinsic contradiction in using models for dating. We will try to discuss this issue in the present paper. Of course, at the same time, the parameters have to be estimated efficiently although there are convergence problems. But, as stated by Pagan (2001), we must "test the model's fit in a number of directions that are meaningful in terms of the objective of the modeling". This can be done by simulating the estimated model and deriving estimates. Regarding turning points, which cannot be among the model parameters, they are not measured easily as stylized facts. Other stylized facts like the phase duration and amplitude may also be an objective for validating the models (Clements and Krolzig, 2000).

For the ex-post dating of turning points, various approaches are available. If an expert committee in the United States is in charge of the dating, Romer and Watson (1994) argued in favor of a systematic method to dating past turning points. The Bry and Boschan (1971) procedure has been used worldwide. Apart from this non-parametric approach, Markov-switching models may be used to dating (Harding and Pagan, 2001, and Krolzig, 2001). Dating seems easy because the past is known by definition. It is true if dating is a concept clearly defined as the peaks and troughs of a series with some censoring rules. This is what has been done for some

[2] The COE, attached to the Paris Chamber of Commerce and Industry, is one of the leading centres in France for economic analysis and forecasting (www.coe.ccip.fr).

decades now in the United States and regularly applied to other countries. The difficulty relates to the identification of the series on which the rule is applied. If GDP were the ideal series, the algorithm could be applied without any question. The only issue would be to make the estimate after clearing seasonality and irregularities and to define minimum censoring criteria (duration, intensity). But it is not the case so that the exercise is done on several coincident series at the same time, which makes the exercise more difficult.

Dating through a parametric model appears quite challenging. Therefore, the finding of Hamilton in its 1989 seminal paper was quite interesting. But the conclusion has been somewhat misleading because there is no reason why this kind of model would, by chance, result in the identification of those turning points. The hypothesis of two regimes is not equivalent to the distinction between recessions and no recessions. And we do not see why this should be the case.

We will give an example of two different economies with different turning points but leading to the same regime-switching model. As a matter of fact, the re-estimation of the Hamilton model later on was quite disappointing and showed that it was not coinciding with official dating anymore. So we will reach the conclusion that this kind of model cannot be used to make a dating of cycles without a clear economic interpretation of the regimes. And therefore, the use of non-parametric methods like the Bry and Boschan (1971) algorithm is the best approach.

We may also think that the definition of recessions might change overtime. Even if recessions have not vanished as some economists were predicting, we may think, if we consider the last US recession, that it has been a mild recession ant that it has not implied a recession in Europe. Only Germany has been facing a short recession. In Europe, we may say that it was a “borderline recession”. As a matter of fact, in many countries, the annual GDP growth rate went down to about 0%, therefore well below the tendencial growth rate.

On the contrary, detection is rather a prediction issue. Real–time estimate of turning points cannot be exact because the identification of a turning point depends on future developments, which are unknown. For example, a censoring rule insists on a minimum duration of 6 months for a recession. This is why the NBER takes so much time to issue its dating. Let us take the example of the recession onset in the USA in April 2002. A real-time (filtered) probability of recession given in that moment by a Markov-Switching model (leading to an emitted recession signal given at that moment by any decision rule) would more likely be interpreted as a prediction. Conversely, the probability estimated by that model computed in November 2002 would not be a prediction but an estimate. The filtered probability includes a component due to the unknown future. Detection models may minimize a clear objective, which may be different from the classical maximum likelihood objective. For example, there may be a need to minimize the number of false signals.

The detection of turning points faces the problem of using data in real time with the difficult issues of edge effects and data revision. It can be thought of as a nowcasting challenge. A turning point may be considered as an event modeled as a binary variable. In this sense, the detection is the probability estimation of the event with an attached decision rule. The detection of leading indicator cycle turning points may be useful to forecast the turning points of the global economy. Markov-switching models are among recent tools for detecting turning points. But the probabilistic approach of Neftçi (1982) is also appealing. Probit and logit models may also be used for that purpose. Finally, pattern recognition procedures have been also proposed. There is quite an exciting development in models aiming at detecting turning points.

Detecting (including anticipating) has a much more intense statistical content because it includes some "predicting" work. In this case, the question is to find the best model to detect as quickly as possible and with the best performance the coming or current turning points. In that respect model-based approaches may be much more powerful than ad-hoc non-parametric rules. For example regime-switching models may be used efficiently to provide efficient signals. But other semi-parametric methods (pattern recognition algorithms or Neftçi's model) may also be powerful. In the case of Markov-switching models, the filtered (real time) probability is more important than the smoothed probability signals because those models are used for prediction purposes. We reach the idea that the best model to predict may not be the best model to explain, and vice-versa.

2.2 Why an official dating is important

Statistical offices have the natural goals of producing and disseminating statistical information among users. This information must meet several well-known criteria, among which: quality, accuracy, timeliness and consistency. In addition, it is now a common practice to disseminate time series data corrected from seasonal variations. First, it prevents the users from doing the correction themselves. Second, it prevents the use of alternative methods, which may result in various estimates while a consensus is more effective. However, alternative methods still exist between countries. The convergence towards a consensual method among countries in Europe, which is encouraged by Eurostat, is a progress towards harmonization.

However, the question of method convergence is gaining complexity when we turn to the issue of trend extraction and characterization of cycles. Even more complex is the issue of identifying turning points in cycles, *i.e.*: dating, detecting or forecasting turning points. In this case, it is needed to clearly define the concepts before reviewing the rather large spectrum of available approaches. There is a growing demand from private as well as from public users for this kind of information to improve decision-making. This is why we will see a multiplication of new tools on that issue, which can be disseminated more quickly through the Internet channel.

Statistical offices may investigate those topics to see if a convergence may be found on a best common approach. But time will be needed, first to identify the concepts and the definitions and second to explore the different techniques available. This will be a step further beyond the correction for seasonal variations or adjustment for trading days towards a better understanding of the evolution of a series around its trend or the comovement of several series around their common trend.

The turning point is a concept, which will have to be clearly defined. Arthur M. Okun (1960), in his famous paper on the appraisal of turning point predictors, had warned, « distinguishing wiggles from genuine turning points is a serious and difficult problem ». The tradition in the last fifteen years has been to distinguish classical cycles from growth cycles. We need to find criteria to characterize turning points of those cycles and avoid focus on short-term lived fluctuations.

Dating cycles finally is geared toward two main objectives: to provide the basis for economic cyclical analysis and to serve as a reference chronology to validate models aiming at "detecting" turning points (Markov-Switching models for example). It could even act as an input in maximization procedures in order to find the best model (Probit models and pattern recognition models for example)

3 Transversal issues

3.1 What cycle and what turning points: The ABCD approach

First, there is a question of definition. In particular, the « business cycle » may either be the « generic » term for economic fluctuations or refer to the fluctuations of the level of the series. In academic literature, this difference is rarely made. In the present study we will distinguish the classical business cycle from the growth cycle (deviation from trend). Refer to the following Figure 1 for the three possible representations: classical cycle (in level), growth cycle (deviation to trend) and growth rate cycle.

The cycle can be defined and estimated in various ways with, for each definition, a different dating of cyclical turning points. It is important to insist on this aspect because the interpretation of turning point signals given by coincident or leading indices will depend crucially on the definition adopted. Unfortunately, there is no unique and recognized terminology to name those cycles. This entails potential confusion for users of indicators available on the market.

We propose a distinction and a terminology based on the dominant tradition that can be met in most empirical works: business cycle and growth cycle. However, it is interesting to distinguish three types of cycles (see Figure 1 for a school case, Figures 2 and 4 for an application on the American economy and Figure 3 for an application on the euro-zone economy). We present in detail these kind of cycles below, as well as our ABCD approach (see also Anas and Ferrara, 2002b). First, the business cycle is meant to reproduce the cycle of the global level of activity of a country. It is the most common definition in the economics literature. The turning points of that cycle (named B for peaks and C for troughs in Figure 1) separate periods of negative growth (recessions) from periods of positive growth (expansions). In real life, a period of negative growth will be admitted as a recession if it respects minimum criteria relative to duration and intensity as well as a sufficient diffusion in the economy (symmetric criteria for expansions). For example, even if growth has become negative in the euro-zone at the end of 2001, this episode did not last long enough to be qualified with certainty as a recession.

The second cycle, largely discussed in Europe, is the growth cycle, introduced by the OECD in the 1960's. This cycle can be defined as the deviation of the reference series (GDP for example) to the trend, though it is difficult to define and estimate. The expression "growth cycle" is not a correct one because it creates confusion with the cycle of the growth rate (discussed below). The growth cycle turning points (named A for peaks and D for troughs in Figure 1) has a clear meaning. Peak A is reached when the growth rate decreases below the trend growth rate. Symmetrically, the trough is reached when the growth rate overpasses it again. In other words, the peak is reached when the first derivative of the deviation to trend becomes null, *i.e.* when the derivative of the series (representing the instantaneous growth rate) equals the derivative of the trend (or the slope of the trend if this one is linear as we assume in the school case presented in Figure 1).

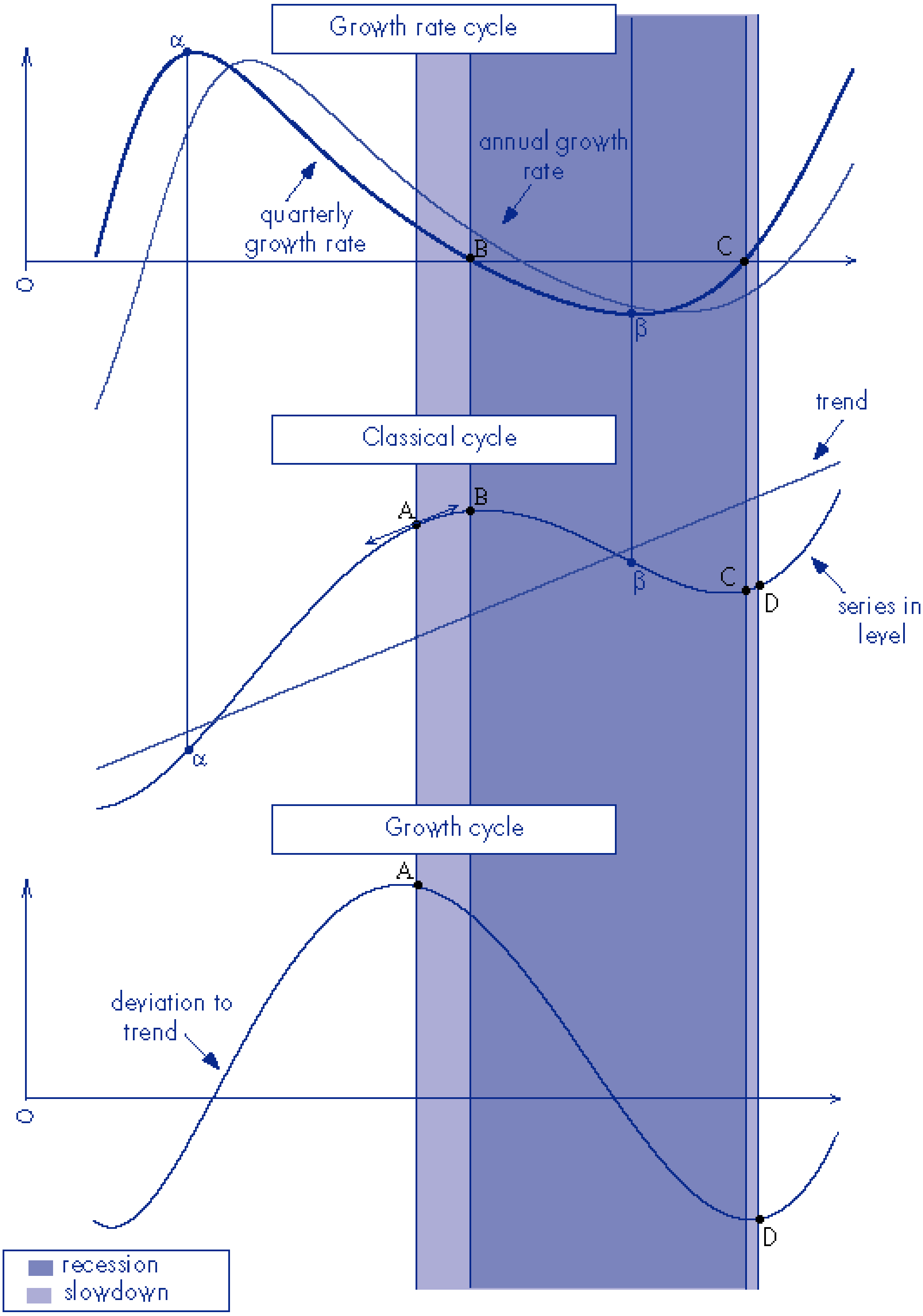

Figure 1: Evolution of cycles and the ABCD approach

Since we need to give a name to those downward and upward phases, we will speak respectively of slowdown and rebound. In French, there is an easy word to refer to the corresponding turning points ("points de retournement conjoncturels") but the translation in English is more complex.

The third cycle we may consider is the growth rate cycle (meaning the cycle of the growth rate). The peak (point α in Figure 1) represents the maximum growth rate. On the contrary the trough (point β in Figure 1) indicates that the growth rate has reached its lower value and is increasing again. It is difficult to give a name to those phases.

Indeed, it is hazardous to speak of slowdown when the growth rate begins to decrease. For example, when the growth rate decreases from 4% to 3% in Europe, it is not a slowdown because growth remains above the trend growth: it is only a slowdown of growth. On the contrary, a renewed increase in growth is not a signal of rebound. Maybe this is what happened in Europe at the beginning of 2002 with a clear upturn of the growth rate but no exit of the growth cycle. On the contrary, it seems that the American economy went out, not only from the recession but also from the growth cycle descending phase.

When the growth rate goes up from –2% to –1%, it corresponds to a phase of decreasing activity (recession) even if the growth rate goes up. Unfortunately, practitioners concentrate most of the time on this cycle. The calculation (on GDP most of the time) is often made on a year-to-year basis and sometimes on a quarterly basis (but more volatile).

However, those calculations have drawbacks relative to the instantaneous growth rate. The year-to-year estimate has a lot of inertia relative to the instantaneous rate and therefore adjust with a delay to the inflexions of growth, notably around peaks and troughs (see Figure 1). If the quarterly change does not have this drawback of inertia, it is nevertheless more volatile, incorporating short-term irregular effects (not seasonal effects because GDP is most of the time corrected from seasonal variations). As regards the underlying instantaneous rate cycle, filtered from short-term perturbations, it is too difficult to estimate.

The ABCD approach considered here is based on the following two principles:

a. The turning points (TP hereafter) detection issue must be considered as the progressive follow-up of the cyclical movement. Instead of concentrating on one TP (a peak for example), it is more informative to consider that the downward movement will first materialize in a peak of the growth rate (point α). Then, if the slowdown gains in intensity, the growth rate will decrease below the trend growth rate (point A) and finally, if it is really deteriorating, the growth rate will become negative (point B) provoking a recession. We get a non-symmetric chronology for upward phases: first exit of the growth rate cycle (point β), second exit of recession (point C) and finally of the growth cycle (point D)

b. We consider that the cycle in growth rates is not a good indicator of future economic cycles. First, it is subject to erratic movements as well as to very short-life fluctuations due to transitory events producing false alarms and making the peak lead extremely unstable, which removes any practical interest for the signal. This is why we consider that the detection of α and β is not always useful nor informative, even if practitioners, market economists or officials often use it for their diagnosis. We prefer to detect points A and B, which announce respectively downward phases of growth cycle and classical cycle.

> However, if the slowdown does not gain in intensity to become a recession, then point A will not be followed by point B. We call the follow-up of those points (A and B for peaks and C and D for troughs) the ABCD strategy for TPs analysis. We will concentrate on the dating, detection and prediction of those TPs. Let's recall that α (growth rate cycle peak) is not necessarily followed by A (growth cycle peak) and could be quickly followed by β. In this case, it creates a weak fluctuation, which does not provoke a sufficient decrease of growth to go below the trend growth rate. This was the case in 1998-99 in the USA during the Asian crisis when the industrial sector underwent a small recession but which did not get diffused to the whole economy, partly due to the reduction in international commodities prices. This is what also happened in France where economists spoke of an "air pocket" (*trou d'air*). Equally, a slowdown will not translate automatically into a recession. The USA experienced a slowdown without recession in 1994-95 but the slowdown starting in May 2000 translated into a recession 10 months later, in early 2001. Meanwhile in Europe, the slowdown did not translate into a recession.

From an economic point of view, it is important to rapidly detect the most pertinent turning points, in order to help public decision making as well as private actor investment decisions. There is a risk of autorealisation process or overreactions.

Nowadays, the Internet channel allows for quick and complete information among users and it is going to be very difficult to slow this process. The multiplication of signals and indicators may or may not be efficient to help users. They will have to decide which indicators to choose or to find the market consensus.

Early signals may be useful to policy makers to take steps sufficiently in advance to avoid the worsening of the economic situation because it takes some time before those measures have an economic impact (3 to 4 quarters for monetary decisions).

3.2 What exactly is detection?

The concept of detection is not commonly used in statistics or in economics. It relates etymologically to the research of an object or phenomena which is "hidden". The detection relates to the research or estimation of a "hidden" event. Under this definition the detection of a turning point is strictly the research and identification of a turning point which has just occurred or is happening in the present time.

However, a wider definition may include, besides the turning point identification in real time, the ex-post dating since, in the past, turning points, even not "hidden", are not clearly observable and need to be estimated. Also, a more general definition could include the detection of coming turning points, *i.e.* the predictions of turning points in the short term. We adopt below this wide vision of the detection issue of turning points.

Let us review the specific issues when identifying successively past, present and future turning points.

There is always the double issue of selecting the adequate series (one or more) used to detect and the selection of the appropriate method (respectively univariate or multivariate). The global performance will depend on both choices.

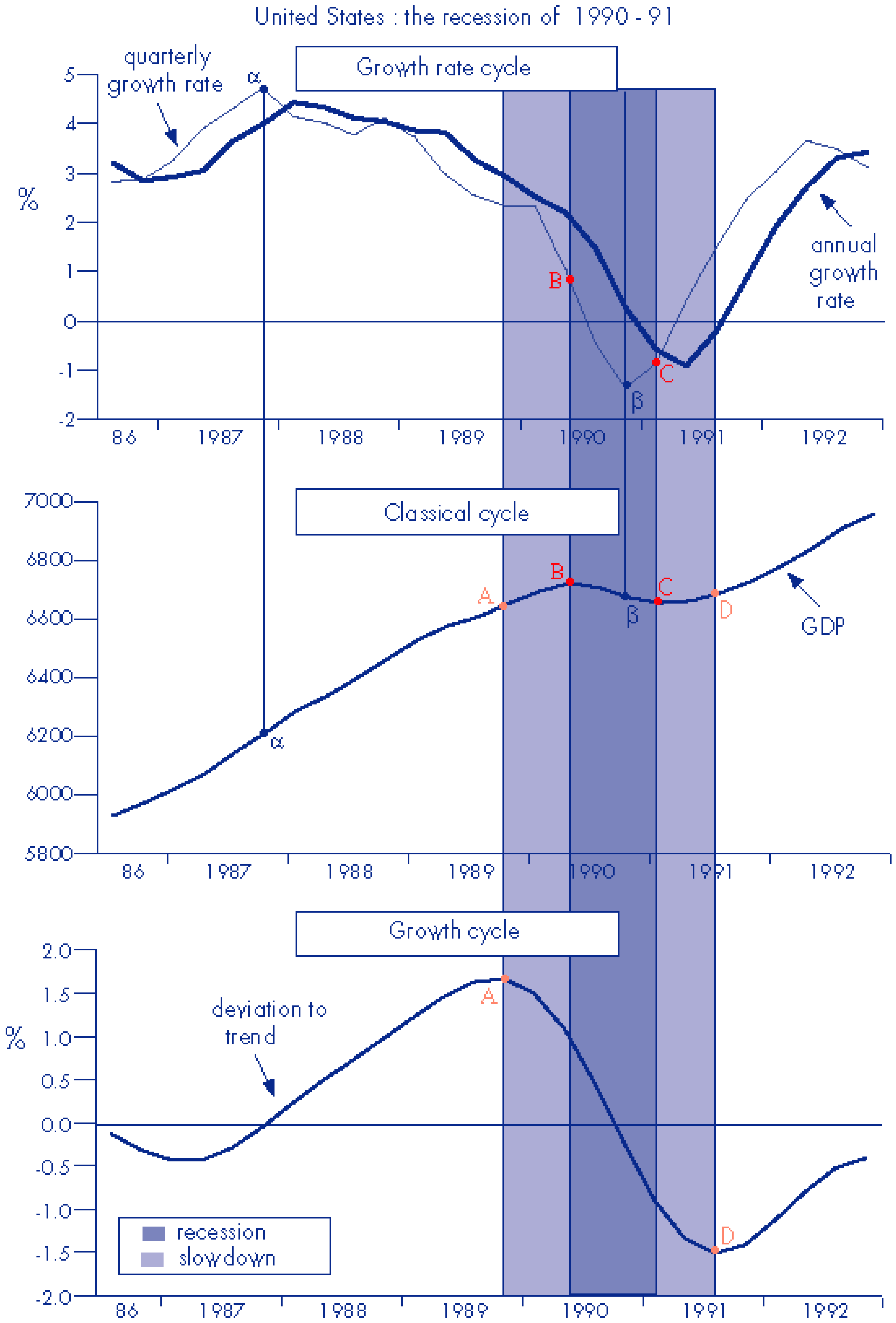

Figure 2: Evolution of the US cycle over the 1986-1992 period and the ABCD approach

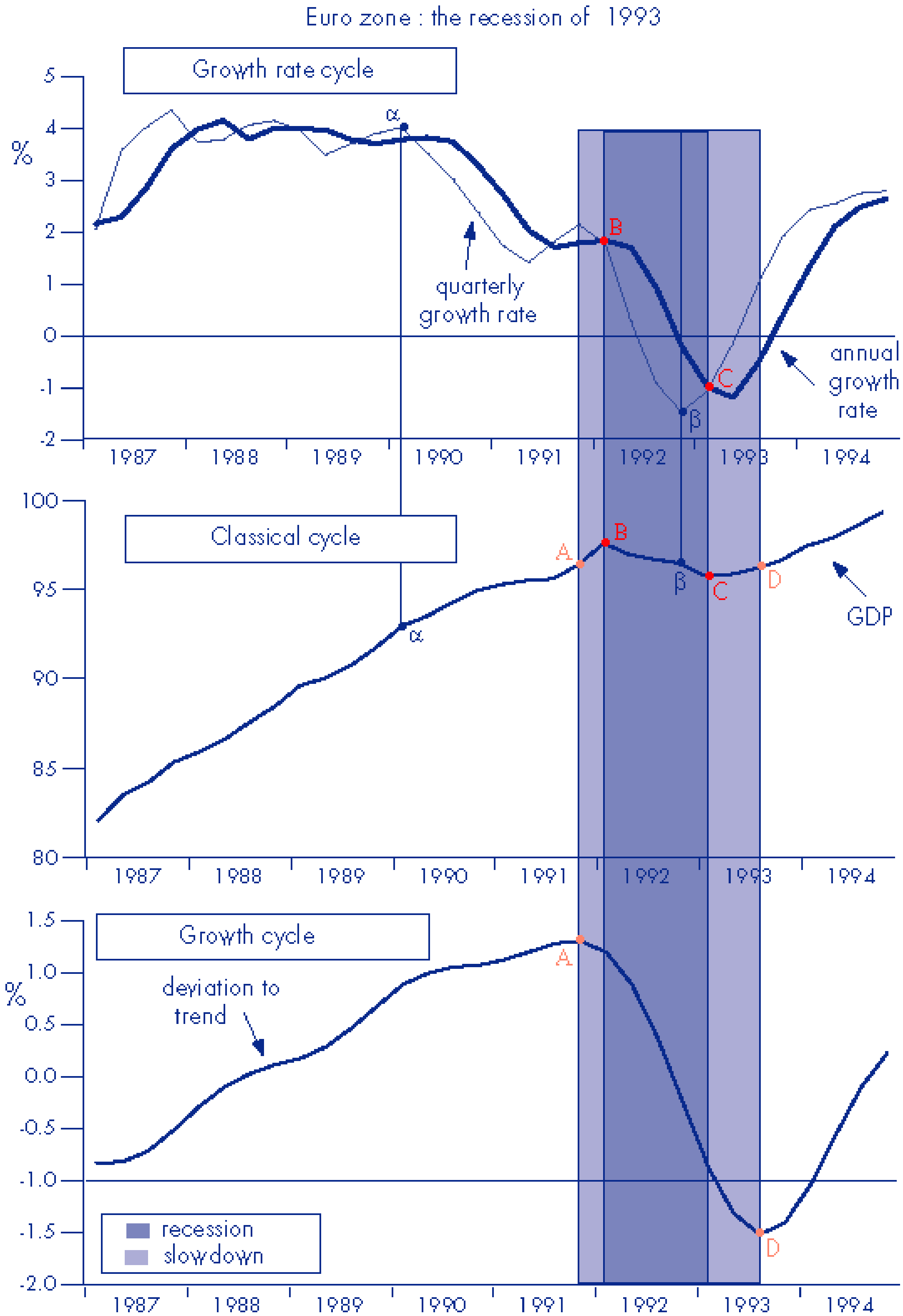

Figure 3: Evolution of the euro-zone cycle over the 1987-1994 period and the ABCD approach

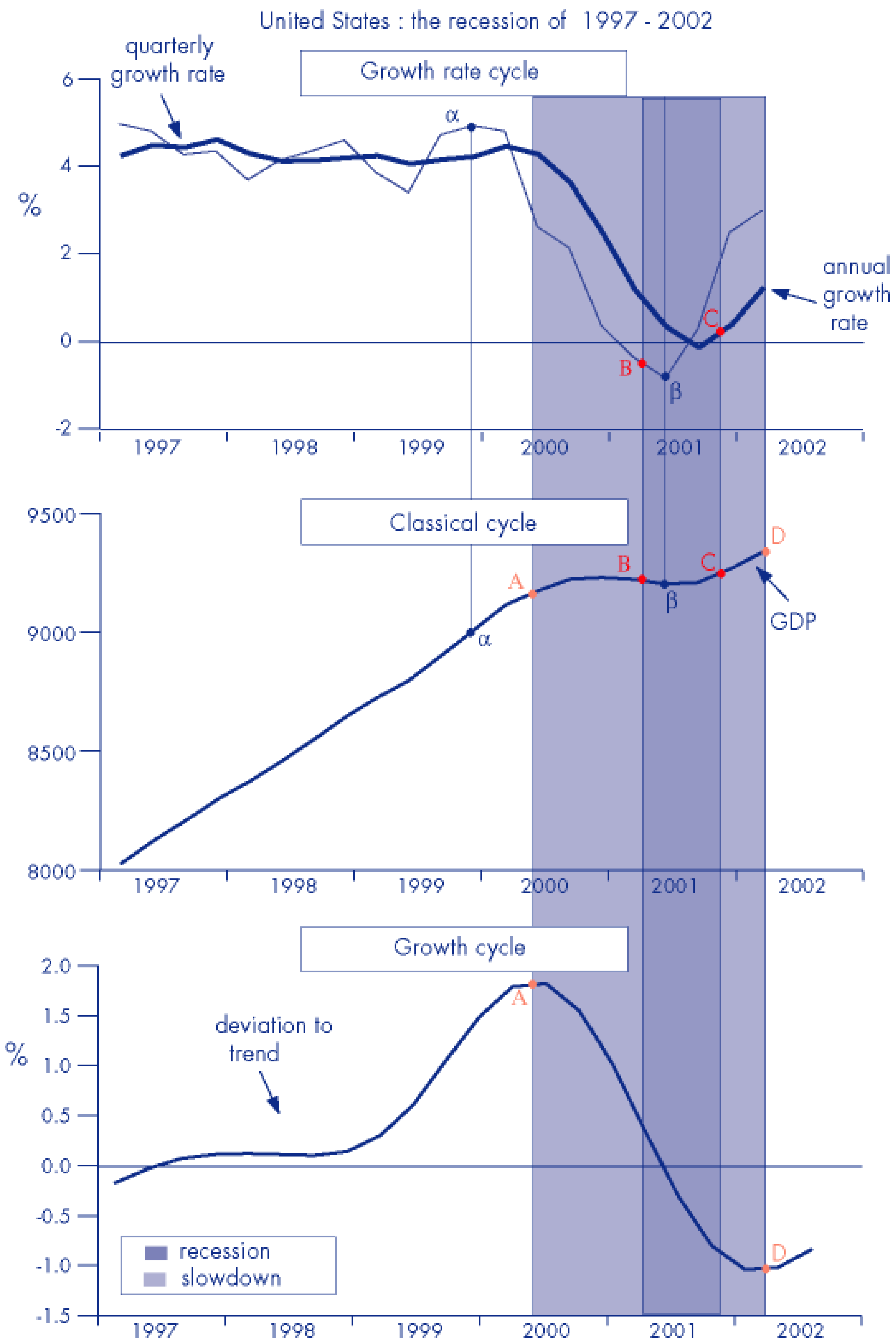

Figure 4: Evolution of the US cycle over the 1997-2002 period and the ABCD approach

3.2.1 Dating the cycles

The detection of past turning points is not easy. In the United States, the NBER's Business Cycle Dating Bureau's Committee is widely recognized as the authority for determining the peaks and troughs of the classical business cycle. There is a substantial delay, however, before the announcement of those dates. For example, the July 1990 peak was announced in April 1991 and the March 1991 trough only in December 1992. It seems however that the NBER business cycle dating procedure has not been homogenous over time. The Economic Cyclical Research Institute (ECRI), a private body in the USA created by G. Moore, is also proposing a dating chronology for a large number of countries. But those dating are neither official nor without critics (Artis, 2002, recently re-estimated the dating of the UK Business Cycle based on monthly GDP estimate). The ECRI also proposed a dating of the growth rate cycle, often confused with a dating of the growth cycle.

In other countries there is no official dating of the classical business cycle. The main issue is the definition of criteria used to recognize an economic fluctuation as a cycle. The Conference Board refers to the 3D's rule (diffusion, deepness, and duration). The dating process, once completed, should be definitive. Therefore, there is a need to avoid absolutely cases of false signals (observe that, on the contrary, it may happen in "real-time detection"). To be sure it is not the case, some delay is necessary to see whether the basic criteria for cycle identification are fulfilled (3 D's for example). But, above all, it is necessary that the data used for dating are definitive, *i.e.* will not be revised any more. This is one of the striking reasons why dating takes some time.

If dating the classical business cycle is not so easy, then dating the growth cycle is even more difficult since the series must first be de-trended. Moreover, the way the series is seasonally adjusted (directly or indirectly for geographic aggregates like euro-zone indicators) and previously adjusted for calendar effects may impact on the dating (see for example Astolfi, Ladiray and Mazzi, 2001, and Lommatzsh and Stephan, 2001). It may therefore happen that different estimates are available on the market.

The objective of dating is more academic that policy-oriented because of the long delay needed accomplish the dating. However, this dating provides an official reference chronology, which allows the measurement of cyclical parameters useful for economic analysis. It above all allows for detection model testing or econometric cyclical modeling validation.

3.2.2 Real time detection

If the use of quarterly GDP series may be sufficient to provide a dating of the past turning points (but this is denied by the NBER Dating Committee), it is clearly not operational in real time. GDP is only available on a quarterly basis with a delay of one to three months, sometimes with significant revisions. GDP is not a good candidate to assess turning points in real time. Thus, the use of other series is unavoidable. Alternatively, one can use a GDP proxy commonly called a coincident index (estimated by use of diverse linear methods). Stock and Watson (1989) revived consideration on co-movement of variables along the cycle by introducing a dynamic factor model in order to extract a common factor. In this case, methods have to be determined to estimate the probability of a turning point of this common factor. In this respect, Diebold and Rudebush (1996) recently proposed to mix together dynamic factor models and regime-switching (see also Kim and Nelson, 1998). If no coincident indices are used, other generally non-linear methods may directly produce the probability of a turning point, like, for example, the multivariate extension of the Hamilton Markov-Switching regime model proposed by Krolzig (1997) and the Markov-switching approach combined with the COE probability aggregation procedure (Anas and Ferrara, 2002a).

An obvious problem is the degree of data revisions. But other issues must be addressed. The edge effect is crucial as the future helps to understand current changes in trends or regimes. There may be two categories of edge effects, the one derived when detrending series by using filters and a *pseudo edge effect* in the use of Markov-Switching models (filtered probability versus smoothed probability). For example, in real time a filtered probability may overpass 50% and send a signal of regime change, while later, the dating based on smoothed probabilities will change the date.

3.2.3 Short-term prediction of a turning point

The timing of the prediction is very important. It is quite difficult to predict turning points in the medium or long term (over 9 months). Even if economic imbalances sometimes make an adjustment plausible or necessary in the future, it is difficult or even impossible to predict when this adjustment will occur. This question is related to the duration dependence of recession's issue and there is no clear definite answer as yet. As a matter of fact, this kind of predictions must take account of exogenous variables, sometimes impossible to predict (external shocks like the September 11, 2001 terrorist attack on New York and Washington for example) and also of future economic policies. In this respect, macroeconomic models are able to design various scenarios in the future based on different sets of assumptions on exogenous variables and policies around the "most probable" scenario.

In the short-term, however, the turning point prediction is or should be easier because, except for important and sudden external shocks, foreseeable changes in economic policies should not reverse the course of economic development due to the impact delay of these measures and the inertia of economic evolution. This is why there is a trend today to complement macroeconomic modeling with short-term leading or coincident indicators. This is done by early detection (prediction) or real time detection. The importance for economic policy is obvious: when slowdowns are predicted, recessions could be avoided or smoothed out by adequate economic policy (but others may think those recessions should not be avoided). In addition, it allows economic private agents to effectively adjust their consumption or investment (but others think it may aggravate the cyclical change or create self-fulfilling expectations), all the more so that agents are not quite rational, according to the last Nobel price in economics.

It is quite difficult to present a classification of all available methods to predict or detect in real time the cyclical turning points. Anticipation of turning points may be possible either because of the existence of leading indicators whose specific turning points precede the global turning points we try to forecast or because the turning point is happening after a progressive slowdown of the series. In other words, the first or second derivatives may be used for anticipating turning points, but unfortunately at the risk of a false signal. In the latter case, we are close to a real-time detection to identify as soon as possible a coming or current global turning point with a minimum risk of error.

But the common feature of all those methods is the need to validate a reference chronology.

Two different approaches are possible. First, find a model and measure ex-post the performance, preferably in a dynamic way, by comparing the resulting predictions with the chronology. This is done with Probit and Markov-switching models. Secondly, include a maximization objective in the estimation of the model directly linked to the purpose: predicting efficiently with the minimum number of false signals. In the second category, the models are less explicative and do not have any purpose of analytical explanation but aims at predicting in the most efficient way.

The approach may be univariate or multivariate. In the univariate case, the model (Probit, Nefçi (1982), Hamilton (1989)) may be applied to a coincident index or to an estimate of the comovement. In the multivariate case, there may be an estimate of the common regime and the co-movement (for example, VAR models in the Stock and Watson (1989) approach, co-movement in dynamic factor models) or an aggregation of probabilities weighted by first type and second type errors (Anas, 1997, Anas and Nguiffo-Boyom, 2001). All those models are competing and it should be possible to compare their performance if their output is clearly understandable. Already, real-time results on turning points prediction are available on some web sites, such as James Stock's (http://ksghome.harvard.edu/~.JStock.Academic.Ksg/), Marcelle Chauvet's (http://faculty.ucr.edu/~chauvet/mc.htm) and COE's one (http://www.coe.ccip.fr).

It is worth mention the comparison made by Chen (2002) between the ad-hoc Conference rule and the use of recession probabilities generated by a Markov-switching model. He started his study on the observation that the 3 D's rule suggested by the Conference Board missed the most recent peak in 2001.

3.3 Detecting with comovement extraction

The "Business Cycle Dating Committee" in USA stated "the Committee gives relatively little weight to real GDP because it is only measured quarterly and it is subject to continuing large revisions". This is why it made an "expert judgment" based on four series.

Many attempts have been made to extract the co-movement of representative coincident series. The co-movement was already stated in Burns and Mitchell's pioneer work (1946). For that purpose, coincident and leading indices were calculated (OECD, Conference Board). More recently, the co-movement has been estimated through factor analysis, either static or dynamic. The method was first applied on a limited number of series (small data mode) and thereafter on big data sets (Watson, 2000, or Altissimo *et al.*, 2001, leading to the EuroCoin project). But, a method will still have to be used to identify the turning point of this series. Either a non-parametric method with an ad-hoc rule (Okun rule, Conference Board rule) or an algorithm (Bry and Boschan), or by inducing probabilities through a Markov-switching approach (the two-step analysis proposed by Diebold and Rudebush, 1996).

Another interesting approach is the direct multivariate approach. In a non-parametric scheme, it is difficult to combine the signals of different series. This is what the NBER Committee tries to do based on expertise. In a parametric scheme, the probability may be calculated by aggregating the signal probabilities of the different indicators (COE approach using first risk and second risk estimates) or by using a multivariate model estimating directly the common regime probabilities (MSVAR of Krolzig, 1997, or Kim and Nelson model, 1998). The application of this approach to big datasets in order to simultaneously extract the co-movement and the regime-switching process will be a tedious effort ahead of us. We will not discuss here if these models can be used for dating or detecting.

Lastly, there is a category of methods clearly designed to detect TP's (predict or detect in real time) and may have a multivariate form, they are considered as pattern recognition models (Oller, 2002 and Keilis-Borok *et al.*, 2000).

4 Methods for dating turning points

There have been many attempts to establish turning point dates by translating the graphical inspection approach into an algorithm. The most famous one is the Bry and Boschan (1971) procedure, still in use in many countries and academic works when estimating business cycle turning points. Apart from these non-parametric approaches, a great number of parametric models have been developed lately to date turning points of the classical business cycle, based mainly on the Markov-Switching model proposed by James Hamilton (1989). An important feature is that all these procedures must be flexible enough to take into account certain characteristics of the cycle such as non-linearity of the process and different durations, amplitudes and cumulative movements of its phases (Harding and Pagan, 2001). We classify these dating procedures in two distinct groups: a) non-parametric procedures; b) parametric procedures.

We will keep this distinction throughout this paper and will now describe both approaches precisely, as well as their major advantages and drawbacks.

4.1 Non-parametric procedures

The first non-parametric procedure consists in examining the relevant time series to locate the peaks and troughs visually (graphical approach). Although not sufficient, this naive procedure can sometimes lead to fruitful results and can be seen as a primary filter.

4.1.1 Dating algorithm

For the most part, non-parametric procedures in turning point dating are based on recognition pattern algorithms. Such algorithms need to perform the following tasks at least:

1. outliers should be disregarded;
2. irregular movements in the series should be excluded;
3. determination of a potential set of turning points;

4. a procedure for ensuring that peaks and troughs alternate;
5. a set of rules that re-combine the turning points established after steps three and four in order to satisfy pre-determined criteria concerning the duration and amplitude of phases and complete cycles.

The predetermined criteria involved in the fifth step of this previous algorithm are referred to as censoring rules. The main censoring rules concern the duration of phases and cycles. For example, dating the business cycle implies the two following rules:

- a phase must last at least six months;
- a complete cycle must have a minimum duration of fifteen months.

However, it appears that the dating results are very sensitive to the choice of the minimum duration phase, as we will see in section 3. For example, when dating the growth cycle, these two previous rules have to be less restrictive; especially the minimum duration of a phase, which is generally set to 12 months.

Some other censoring rules may also be considered:

- in the presence of a double turning point, the last turning point must be chosen;
- turning points within six months of the beginning or end of the series are disregarded.

Regarding the third step of the algorithm, several ways allow the identification of the potential turning points. We present three of the most frequently used in practice. First, let us note $(y_t)_t$ the time series of interest and adopt the following convention, for all date t: $\Delta y_t = y_t - y_{t-1}$ and for each integer k, $\Delta_k y_t = y_t - y_{t-k}$. The best known approach, widely released in the media to detect real time peaks and troughs in the classical cycle is the following:

$$\begin{array}{ll} \textit{Peak at } t: & \{\Delta y_{t+1} < 0, \Delta y_{t+2} < 0\} \\ \textit{Trough at } t: & \{\Delta y_{t+1} > 0, \Delta y_{t+2} > 0\} \end{array} \tag{1}$$

This rule has been attributed to Arthur Okun by Harding and Pagan (1999). It means that a recession involves at least two quarters of negative growth. This rule is generally applied to the quarterly GDP. Another approach can be found in Wecker (1979) and has been used in Pagan (1997):

$$\begin{array}{ll} \textit{Peak at } t: & \{\Delta y_t > 0, \Delta y_{t+1} < 0, \Delta y_{t+2} < 0\} \\ \textit{Trough at } t: & \{\Delta y_{t-1} < 0, \Delta y_t < 0, \Delta y_{t+1} > 0\} \end{array} \tag{2}$$

This second rule is also generally applied to the quarterly GDP to identify peaks and troughs in the classical cycle. Lastly, the third approach describes the heart of the Bry and Boschan (1971) algorithm and is given by:

$$\begin{array}{lll} \textit{Peak at } t: & \{y_t > y_{t-k}, y_t > y_{t+k}, & k = 1, \ldots, K\} \\ \textit{Trough at } t: & \{y_t < y_{t-k}, y_t < y_{t+k}, & k = 1, \ldots, K\} \end{array} \tag{3}$$

where $K = 2$ for quarterly time series and $K = 5$ for monthly time series. All these three approaches are based on a variation in growth rates over a bandwidth in comparison with an *a priori* threshold set to zero. The choice of the threshold value is somewhat natural in this case.

4.1.2 Multivariate framework

It must be pointed out that this kind of pattern recognition algorithm is well designed to provide a turning points dating chronology when applied to a single time series. However, in economic cycles analysis, an important feature to be considered is the concept of co-movement, pointed out by Burns and Mitchell (1946) in their seminal work. Co-movement refers to the fact that most macroeconomic time series evolve together along the cycle. That is, the practitioner is placed into a multivariate framework. Generally, non-parametric datings are applied to a univariate time series representing the co-movement, such as GDP, industrial production or indexes stemming from dynamic factor analysis (see for example the Euro-COIN indicator). However, these previous proxies are often criticized and their use can sometimes lead to spurious cycles. As an application, in section 3 of this paper, we try to use the industrial production as a proxy of the whole production.

Therefore, practitioners prefer to have a global overview of what is going on by considering a set of series. For instance, the Dating Committee of the NBER considers a small set of four monthly time series (industrial production, employment, real income and wholesale-retail sales) to construct the business cycle chronology of the American economy. In this case, the difficulty lies in the summary of the diverse turning point dates obtained. The aggregation procedure consists more of *ad hoc* rules and experts claims than a rigorous mathematical algorithm. Thus, this class of procedure suffers under a lack of transparency. It is indeed a hard task to reproduce exactly the procedures and the dating results provided by the NBER or the Conference Board.

Another issue linked to the multivariate framework is specific to large economic areas including several national economies. In order to provide a turning point chronology, is it more appropriate to analyse the economies of each country of the zone (indirect approach), or the whole economy of the zone directly (direct approach)?

Regarding the indirect approach, the most difficult part is the way to aggregate the multivariate information. Once we get a turning point chronology for each country of the euro-zone, it is necessary to define a way to summarise this information to provide a chronology for the euro-zone. Some criteria have been proposed in literature. For instance, Krolzig and Toro (2001) argue that Europe is considered to be in recession if at least half of the countries are in recession. However, it seems that this criterion is rather arbitrary and is not based on economic rationale. Another approach would be to weight the information of each country by a measure representing the importance of an economy in the whole euro-zone.

Consequently, it is worth noting that in the multivariate framework, non-parametric procedures are very difficult to adapt.

4.2 Parametric procedures

Parametric dating procedures are based on statistical modeling of a given set of selected time series. Therefore, if we consider a single time series (GDP, IPI or any coincident index), it involves univariate modeling and if we consider more than a single series, it involves multivariate modeling. In this section, we consider only hidden Markov-chain models, as for instance the model proposed by James Hamilton (1989). Although other time series parametric models are sometimes used (Logit and Probit models, for example), the most often used model is the Markov-Switching regime.

4.2.1 The Markov-Switching model

The Markov-switching regime model (MS hereafter) has been introduced by James Hamilton (1989) in order to take into account a certain type of non-stationarity inherent to some economic or financial time series that cannot be caught by classical linear models. Having observed that such time series frequently exhibit shifts in mean, Hamilton's original idea was to model these non-stationary time series by using a piecewise stationary linear process. It is often assumed that the observed time series are approximated by an autoregressive process, whose parameters evolve through time. Moreover, their evolution is ruled by an unobservable variable which in turn follows a first order K-state Markov chain, independent of past observations on the observed time series.

Markov-switching models have been extensively used in practice, we refer, for example, to Anas and Ferrara (2002a) for a review of applications in the macroeconomic field and to Franses and van Dijk (2001) for a review in applications in the financial field. Regarding the theoretical aspects of the model, we refer mainly to the monography of Hamilton (1994) and to Krolzig (1997) for multivariate generalizations.

We recall the main definitions related to the MS model. The Markov-switching process of Hamilton, $(Y_t)_t$, in the case of an VAR(p) process, is given by the following equations:

$$Y_t - \mu(S_t) = \sum_{j=1}^{p} \phi_j(S_t)(Y_{t-j} - \mu(S_{t-j})) + \varepsilon_t \tag{4}$$

where $(\varepsilon_t)_t$ is a Gaussian white noise process with finite variance $\sigma^2(S_t)$ and where the unobservable variable $S_t = 1, .., K$ is supposed to represent the current state of the economy. This previous form is known as the mean adjusted form of the MS(K)-VAR(p) model. Another well known, but different, specification of the model is the intercept form and we refer to Krolzig (1997) for a description.

Moreover, the whole specification of Markov-switching model needs the specification of $(S_t)_t$ as a K-state first order Markov chain. That is, the value of the time series S_t, for all t, depends only on the last value S_{t-1}, *i.e.*, for $i, j = 1, \ldots, K$:

$$P(S_t = j | S_{t-1} = i, S_{t-2} = i, \ldots) = P(S_t = j | S_{t-1} = i) = p_{ij}. \tag{5}$$

The probabilities $(p_{ij})_{i,j=1,\ldots,K}$ are called *transition probabilities* of moving from one state to the other. Obviously, we get for $i, j = 1, \ldots, K$:

$$\sum_{j=1}^{p} p_{ij} = 1. \tag{6}$$

4.2.2 Estimation

We are now interested in parameter estimation of a MS(K)-VAR(p) process and in the estimation of the probability of being in a given regime. Let $(y_1, \ldots, y_T)_t$ be an observed time series with finite sample size T generated by a MS(K)-VAR(p) process. We assume that the parameter θ to be estimated. The parameter estimation method generally used is the classical maximum likelihood estimation (MLE hereafter) method, based on the assumption that the white noise process $(\varepsilon_t)_t$ in equation (4) is a Gaussian process. Furthermore, the Markov chain $(S_t)_t$ is supposed to be independent of $\varepsilon_{t'}$, for all t and t'. The MLE method is somewhat classical in the statistical literature, but in this case the main difficulty stems from the fact that the latent process S_t cannot be observed and has therefore to be estimated, for all dates t. The MLE method aims at finding the parameter θ so that the conditional log-likelihood $L(\theta)$ is maximum, with $L(\theta)$ expressed as :

$$L(\theta) = \sum_{j=1}^{p} \log f\left(x_t / F_{t-1}, \theta\right), \tag{7}$$

where, for all t, F_t denotes the vector of observations obtained through date t and where $f(x_t/F_{t-1}, \theta)$ is the conditional density of the MS(K)-VAR(p) model.

Some practical experiences have shown that the likelihood function of an MS model is ill behaved, exhibiting some local maxima. Therefore, the choice of the starting values of the algorithm is crucial to find the global maximum. Some empirical studies carry out ML estimation of the parameters by implementing the Expectation-Maximisation (EM) algorithm proposed by Hamilton (1990). This algorithm has been proved to be more robust to the initial parameters values.

Thus, the log-likelihood $L(\theta)$ can be evaluated for a given parameter θ. However, the evaluation of the conditional log-likelihood $L(\theta)$ asks for the knowledge of $P(S_t = i/F_{t-1}, \theta)$, for $i = 1, 2$. This estimation is computed by using properties inherent to Markov chains: this is the forecast of being in the state i given the information through date $t-1$. This estimated probability $P(S_t = i/F_{t-1}, \theta)$, for $i = 1, 2$, is referred to as the filtered probability of being in state i. The conditional probability of being in the state i can be computed, given all the available information through date T. This probability $P(S_t = i/F_T, \theta)$, for $i = 1, .., K$, is referred to as the smoothed probability of being in state i.

4.2.3 Dating and validation

Once one has estimated the parameters from the data, we get as a by-product the filtered probability $P(S_t = i/F_t, \theta)$ and the smoothed probability $P(S_t = i/F_T, \theta)$ at every time t. This latter probability is then used to classify the observations between regimes. A simple rule is to assign the observation at time t to the regime with the highest smoothed probability:

$$i^* = ArgmaxP(S_t = i/F_T, \theta). \tag{8}$$

For the simplest case of two regimes, the rule consists in assigning the observation to the regime i with $P(S_t = i/F_T, \theta) > 0.5$. It is also possible to date turning points following this algorithm:

$$\begin{array}{ll} \textit{Peak at } t: & \{P(S_{t+1} = 1|y_1, \ldots, y_{T^*}) < 0.5, P(S_t = 1|y_1, \ldots, y_{T^*}) > 0.5\} \\ \textit{Trough at } t: & \{P(S_{t+1} = 1|y_1, \ldots, y_{T^*}) > 0.5, P(S_t = 1|y_1, \ldots, y_{T^*}) < 0.5\} \end{array}$$

This algorithm simply means that a turning point is a change in regime. It is worth noting that the choice of the threshold 0.5 is "natural" and may lead to controversy.

However, a validation stage is needed to test the ability of the model to replicate business cycle characteristics (Krolzig and Toro, 1999 and Harding and Pagan, 2001). Indeed, the regime estimates, and therefore the implied dating, do not have spontaneously a clear interpretation. *A priori*, there is no reason why the estimated regimes would coincide with the economic cycles phases. To illustrate this idea, we first have estimated a Markov-switching model on the US GDP quarterly growth rate, over the period 1948-Q3 to 2002-Q3. The model is quite simple, without any autoregressive terms and with a constant variance across both regimes (model 1). We applied the same model to the same series increased uniformly by 1% (model 2). As can be seen in the Figure 5, the periods of negative growth rates are obviously more seldom. Therefore, if we refer to the classical definition of recessions, the number of recessions in both series is different. For example, in the second case, there is no recession in 1990-91, nor in 2001. When we estimate the models, we clearly see that the estimated smoothed probabilities are equal. As far as the parameters are concerned, only the regimes average is modified, as can be seen in Table 1. The growth rate of high regime (expansion) is increased by 1% from 1.2% to 2.2% (see Table 1) while the negative growth rate of recession (-0.1%) becomes positive (0.9%).

Table 1: MS(2) model on US GDP: estimated coefficients

	Model 1	Model 2
μ_1	-0.1%	0.9%
μ_2	1.2%	2.2%
p_{11}	0.77	0.77
p_{22}	0.92	0.92
σ	0.85	0.85

Note: Growth rate (model 1) versus growth rate raised by 1% (model 2)

This previous example demonstrates that the Markov-switching model is not able to distinguish periods of recessions, as defined by common tradition. It only separates regimes in accordance to the specification of the model. This separation will be different if we change the specification: variances depending on regimes, time-varying transition probabilities, autoregressive terms, etc.....

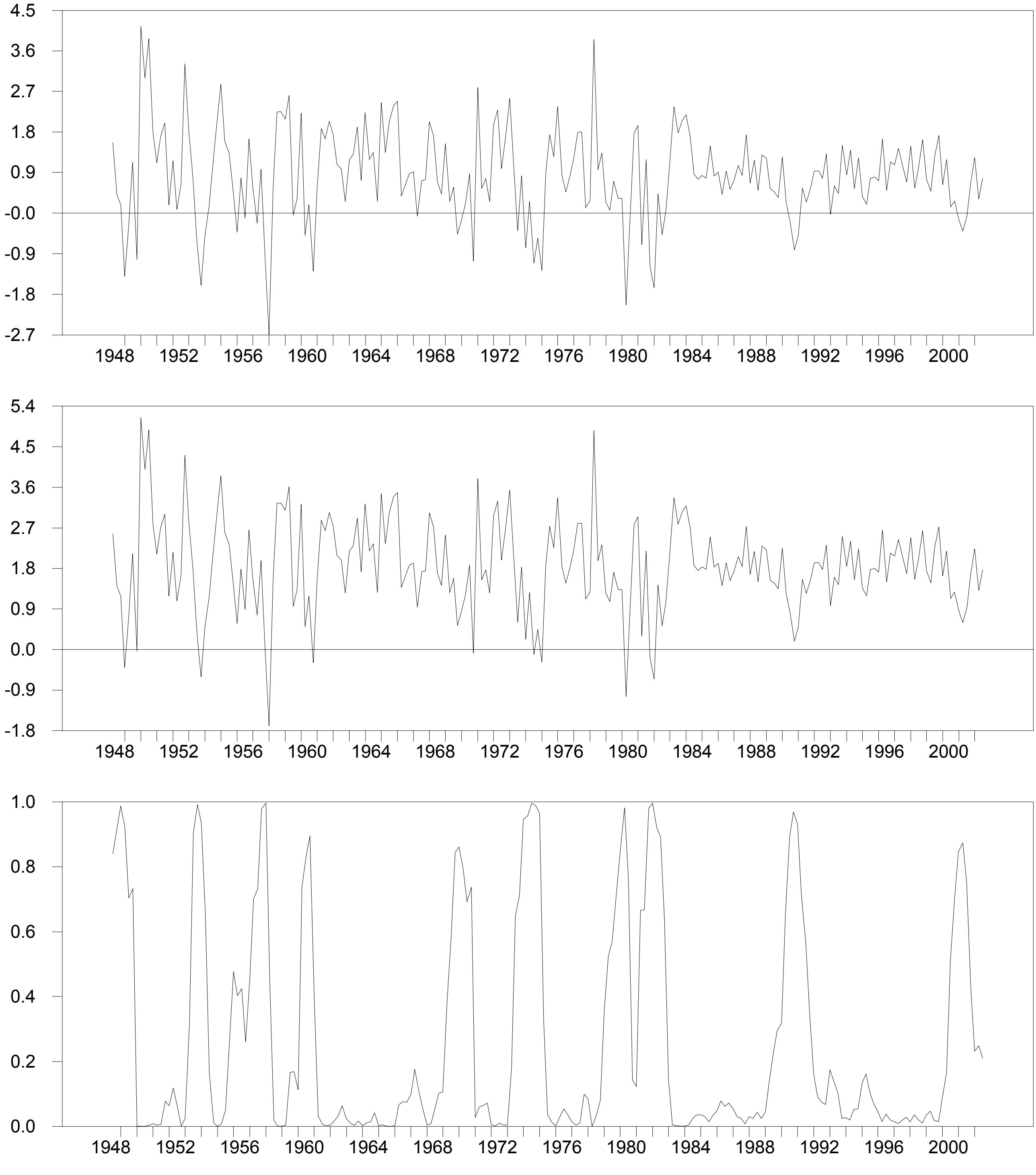

Figure 5: US GDP growth rate versus growth rate raised by 1%, from 1948 –Q3 to 2002-Q3

It is note sure that we may find the best specification by minimising criteria like AIC or BIC, based on the residual variance. On the contrary, many alternative models, *i.e.* representations, are possible. For example, it seems that there are equivalent combinations of estimates of autoregressive terms and transition probabilities. Both parameters try to capture the time dependence of data.

This does not prevent us from finding a plausible specification, which replicates reasonably the official dating in the past. In that case, we may say that the model is well fine-tuned and could help in the future to date the cycle. But what do we gain by doing that? We just replicate a mechanism, which gives good results. There is no clear advantage in doing that. The fact that the official dating could have been replicated by this kind of model is satisfactory because it shows that there probably exists an underlying cyclical process which can be statistically represented. The clear advantage of such model, which usefully replicates dating, is to become a good instrument to detect the cycle in real time.

Another clear illustration of the difficulty in finding the adequate model to replicate the business cycle is found looking for the GDP euro-zone Markov-switching model. Replicating the simple Markov-switching model with one autoregressive lag estimated by Krolzig (2001) over the period 1982-Q2 to 2000-Q4 and extended until 2002-Q2, we find a recession in 2001 (see Figures 6 and 7). This is puzzling since no recession has been observed recently in Europe.

We may underline two phenomena. First, the fragile estimation of the model, the period of estimation only incorporates two recession episodes (one at the beginning of the period may not be well defined) which does not allow for a robust estimation of transition probabilities. In econometric models, there is a need for sufficient observations but in a Markov-switching model, what is important is the number of regimes observed. The second observation is related to the instability of the parameters.

If we begin the model estimation in 1970, the presence of a third recession (linked to the first oil shock) changes the estimation of the parameter and the clearer distinction of recession in terms of negative growth rate imply a "no-recession" in 2001.

It is interesting to see what changes when we consider a third regime, starting the sample data in 1980. If we consider three regimes instead of two, we observe that the model does not capture the 2001 recession (see Figure 8). Those apparently contradictory results seem puzzling, but the results depend clearly on the specification of the model. We need to find the best representative model, that is the best to replicate the past dating.

The use of a third regime is generally made with another objective. For example, Krolzig (2001) mentions that on the period starting in 1970, there in an issue when working in a multivariate approach with several European countries including Southern countries (Spain and Italy particularly) because those countries had much higher growth rates in the 1970's.

Therefore, the third regime is only capturing this "time" regime, or in other words, it is a way to deal with the non-stationarity in mean of the series. Starting in the 1980's, he does not need this third regime anymore to capture this effect. But a third regime may have another objective: to make a better separation between cycles, to help classify the intermediate cycle (the growth cycle) and to estimate the business cycle more clearly. Those examples show the difficulty in finding the adequate specified model to replicate the dating chronology. The objective may be other than replicating the dating, it could be the representation of "stylised facts" of the business cycle as was done for example by Clements and Krolzig (2000).

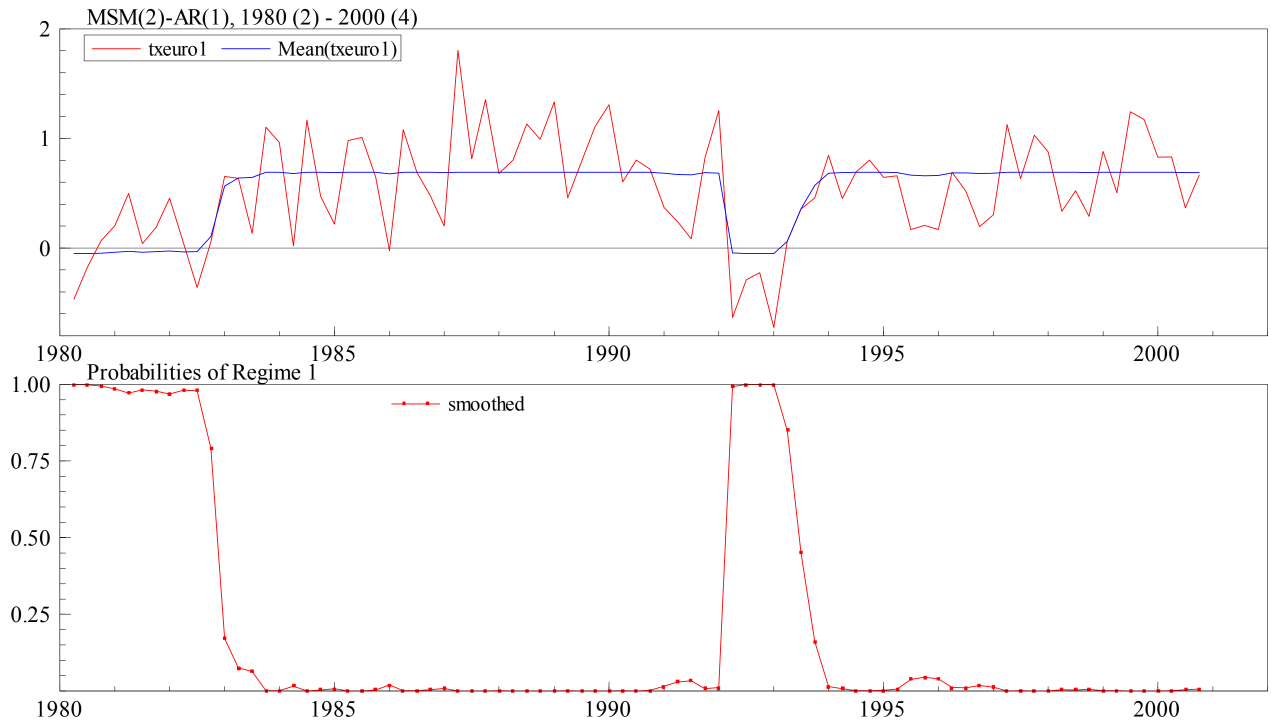

Figure 6: Smoothed probabilities of regimes of an MS(2) model applied to the euro-zone GDP from 1980-Q1 to 2000-Q4

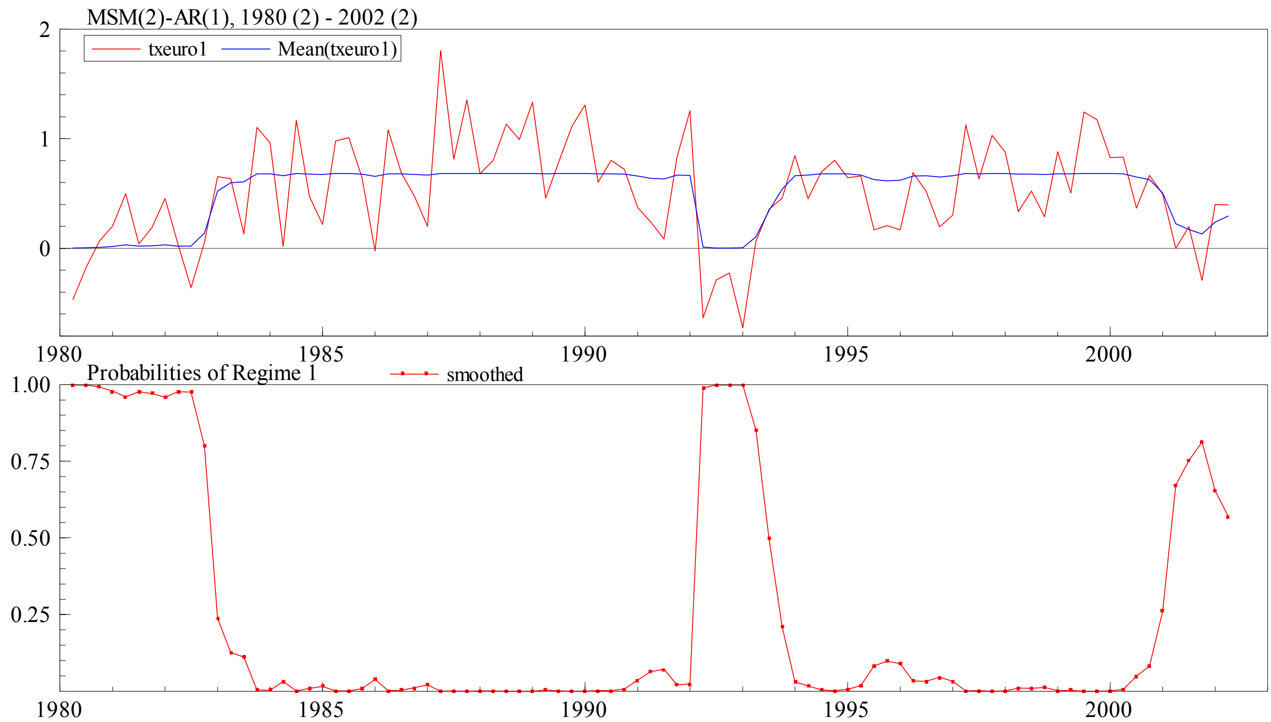

Figure 7: Smoothed probabilities of regimes of an MS(2) model applied to the euro-zone GDP from 1980-Q1 to 2002-Q2

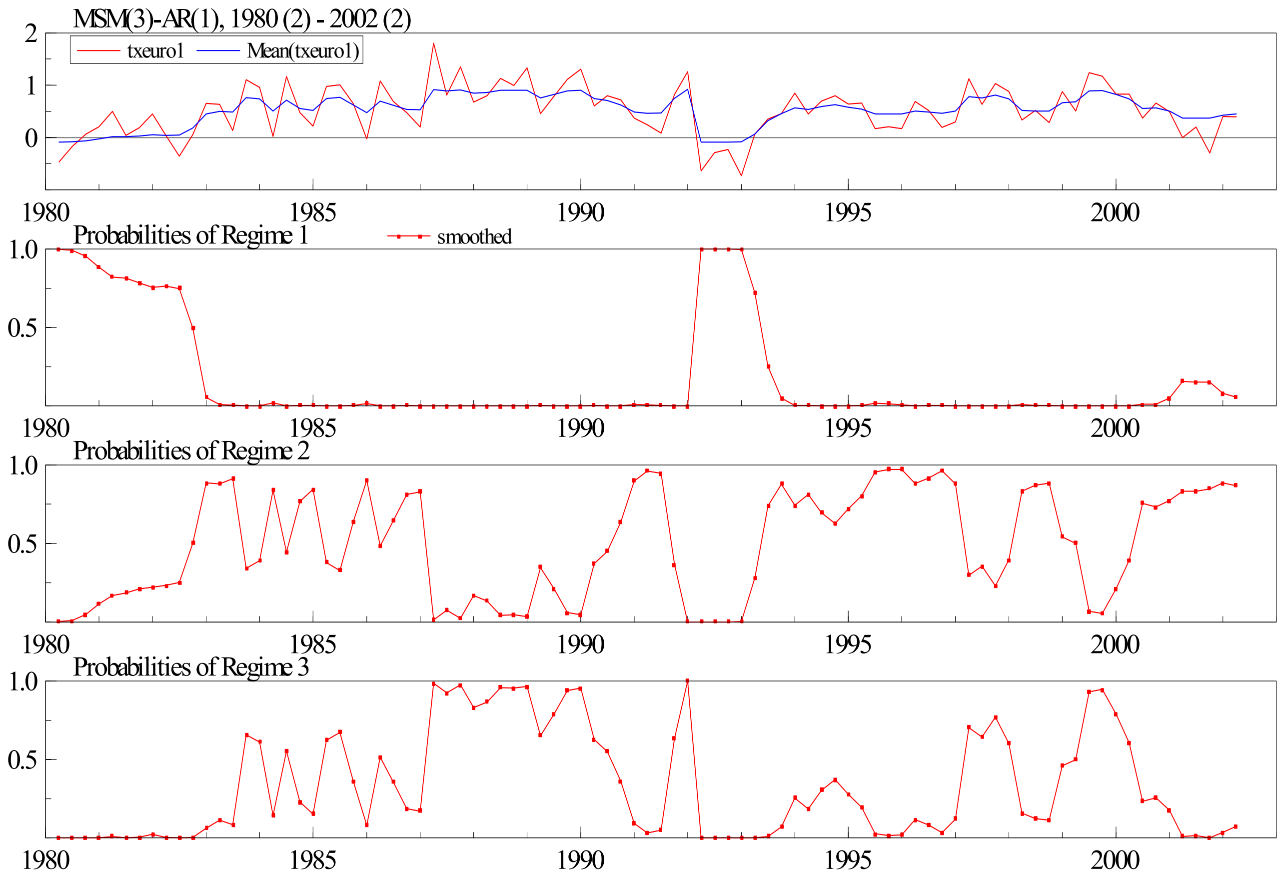

Figure 8: Smoothed probabilities of regimes of an MS(3) model applied to the euro-zone GDP from 1980-Q1 to 2002-Q2

By using numeric simulations on an estimated model, they try to assess if some stylised facts, like phases durations, are respected. Whatever the objective, the idea is to find the best specification to reach the objective. Therefore, there is no reason why an MS model would provide directly the dating of the business cycle. Hamilton model estimated in the 1989 paper was able to replicate the dating. But if we were controversial, we would say that it could have been by chance since there was no optimisation for that purpose. Finally, what happened is that a few years later, when the model was re-estimated, it did not work anymore and the specification had to be changed by adding a third regime or by introducing variable transition probabilities. Therefore, those models are not able to produce an optimal dating but we may find a specification, which helps to reproduce the dating. This is almost a calibration of the model. It is interesting because it gives a theoretical foundation for the existence of pseudo-cycle in series, represented by a stochastic process, argument which have been developed by James Hamilton in an answer to Don Harding[3]. But what is more important is that those models that give a reasonable representation of the business cycle can be used thereafter for prediction or real-time detection of turning points.

[3]See James Hamilton's web page http://www.econ.ucsd.edu/~jhamilto/.

5 Comparison of non-parametric and parametric methods in dating the cycle: application to the case of the euro-zone

In this section, a comparative assessment of non-parametric and parametric methods is carried out on real data. Especially, we compare a version of the Bry and Boschan approach with a Markov-Switching model in order to date the growth and business cycles of the euro-zone industrial production index. The data set used is the monthly euro-zone industrial production index (IPI) from January 1971 to August 2002. The series is seasonally adjusted and is presented in Figure 9. Note that we could have used GDP to describe the co-movement of the whole economy. One of the drawbacks is that GDP is sampled on a quarterly basis; a monthly dating is more accurate (the OECD uses both GDP and the industrial production to produce a monthly reference chronology, see OECD, 2002). Moreover, it is often difficult to get a long historical record of GDPs at the desired frequency and some statistical procedures, such as back-calculations, are needed. The choice of the series is not important insofar as the objective is to compare the methods. It is also assumed that there is a convergence in the cyclical movements of the European countries and we therefore use the European aggregate as a proxy for the co-movement.

5.1 Model specification

We describe precisely both parametric and non-parametric methods developed in this application to date the turning points of the growth and business cycles.

5.1.1 Non parametric procedure

The non-parametric procedure developed in this section is based on the following algorithm:

- the IPI series has not been corrected from outliers;
- irregular movements in the series are excluded by using a 3-terms centred moving average (the dating chronology does not seem to be so sensitive to the degree of smoothing, the results obtained with a 3-terms and a 5-terms moving average are similar);
- a potential set of turning points on the time series of interest $(y_t)_t$ is determined by using the following rule, which is the heart of the Bry and Boschan (1971) algorithm:

$$\begin{array}{ll} \textit{Peak at } t: & \{y_t > y_{t-k}, y_t > y_{t+k}, \quad k = 1, \ldots, K\} \\ \textit{Trough at } t: & \{y_t < y_{t-k}, y_t < y_{t+k}, \quad k = 1, \ldots, K\}, \end{array}$$

  where $K = 5$, as usual for monthly time series.
- a procedure for ensuring that peaks and troughs alternate is developed by using the following rule:

(i) in the presence of a double through, the lowest value is chosen

(ii) in the presence of a double peak, the highest value is chosen

- the censoring rules concerning the duration of phases and cycles are the following:

(iii) a phase of the cycle must last at least six months

(iv) a complete cycle must have a minimum duration of fifteen months

(v) turning points within six months of the beginning or end of the series are disregarded.

Note that we let the same minimum duration phase of six months for both growth and classical cycles.

Concerning the growth cycle turning points chronology, we consider two different de-trending methodologies: the Baxter-King and the Hodrick-Prescott filters (respectively BK and HP hereafter).

Note that the growth cycle extracted from the HP filter has to be smoothed to extract the irregular component while the growth cycle extracted from the BK filter hasn't.

From a technical point of view, we keep cycles with a period lying between 1.5 and 8 years regarding the BK filter, and regarding the HP filter the arbitrary constant λ is set to 300 000.

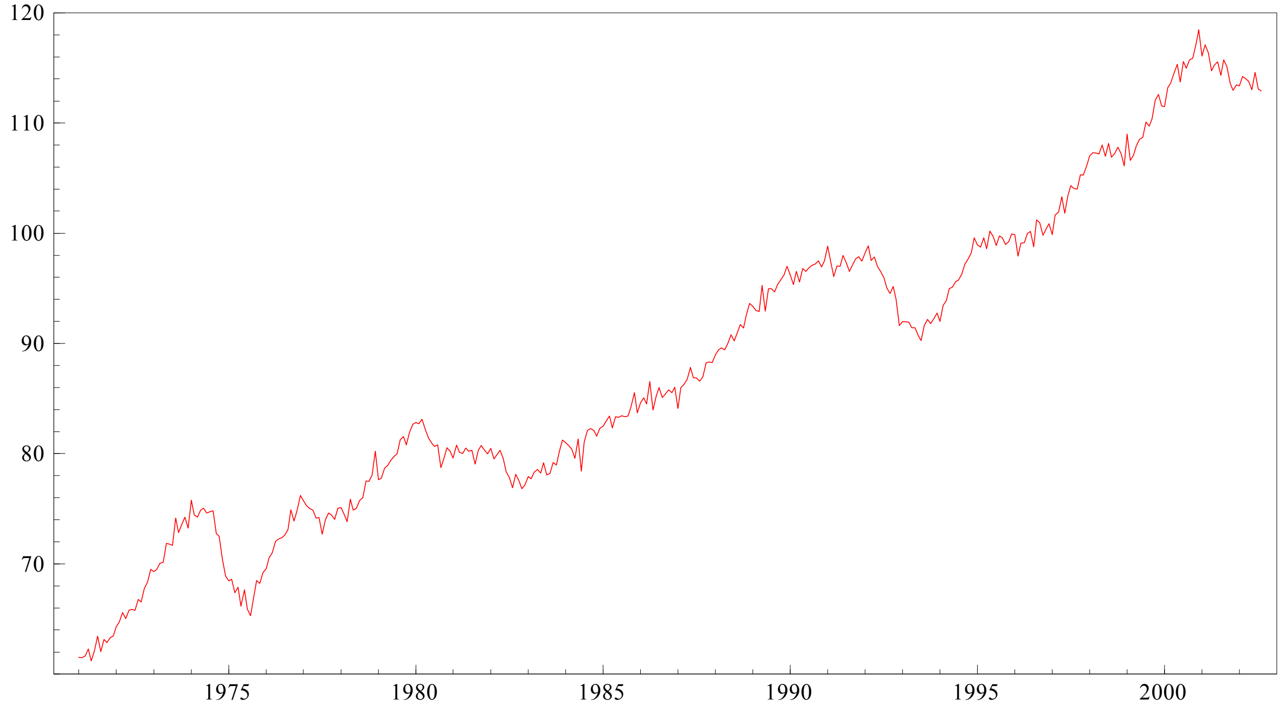

Figure 9: Monthly IPI in the euro-zone 1971-M1 to 2002-M8

5.1.2 Parametric procedure

In order to identify *a posteriori* the turning points of the euro-zone business and growth cycles, a Markov-Switching model is applied to the euro-zone industrial production index growth rate over three months.

In economics, the unobservable variable describing the regimes, denoted $(S_t)_t$, is often supposed to represent the current state of the economy. Thus, a 2-state Markov chain is generally used in applications, that is, for all t, the time series S_t takes value 1 when the economy is in contraction and value 2 when the economy is in expansion. However, some authors found evidence in favour of a three-regime model for the business cycle (Sichel (1994), Boldin (1996), Clements and Krolzig (1998), Krolzig and Toro (1999) and Layton and Smith (2000)). Either the expansion phase can be separated into a regular growth phase and a high growth phase, or the contraction phase can also be separated into a slowdown phase and a recession phase. Therefore, the number of states must be discussed before the modelling and asks for a statistical test. Unfortunately, as noted by Hamilton (1994, p. 698), the assumption that the process describing the data presents a given number K of regimes cannot be tested using the usual likelihood ratio test, because some of the classical regularity conditions are not fulfilled. The most efficient specification test is the one proposed by Hamilton (1996), who takes the $(K-1)$-state model as the null hypothesis and conducts a variety of tests to check whether a K-state model is needed.

Our aim is to develop a single model for dating simultaneously the growth and classical cycles. Therefore, our choice is to consider a three-regime Markov-Switching model (MS(3)) with the hope that each regime will have the following economic interpretation :

- low regime ($S_t = 1$): the economy is in recession (low phase of the classical cycle)
- intermediate regime ($S_t = 2$): the growth rate of the economy is below its tendencial growth rate (low phase of the growth cycle without recession)
- high regime ($S_t = 3$): the growth rate of the economy is over its tendencial growth rate (high phase of the growth cycle)

Note that the present interpretation of the three regimes implies a constant long-term growth rate over the whole sample period.

The chosen MS model is specified as having a different variance for each regime and no autoregressive lag. Regarding the autoregressive order p, the original Hamilton $p = 4$ choice was often criticised in the literature, because the autoregressive parameters are proved to be frequently significantly equal to zero. Even if they are significant, Lahiri and Wang (1994) have shown that imposing any degree of autoregression in the errors on the basic MS model deteriorates the performances in terms of turning point detection. Lastly, we try some different values of p, $p = 1, 2, 3, 4$, and it turns out that the choice $p > 0$ strongly deteriorates the interpretability of the model.

Parameters are estimated by maximum likelihood estimation trough the expectation-maximisation (EM) algorithm, as in the paper of Hamilton (1990). Practically, the estimation is carried out by using the MSVAR module for Ox (see Doornik (1999)) developed by H. M. Krolzig . Estimated parameters of the model are presented in Table 2.

In order to provide a dating chronology, we get the three smoothed probabilities of being in each regime, denoted $P(S_t = i|y_1, \ldots, y_T)$, for $i = 1, 2, 3$, to classify the observations between regimes (see Figure 10). Regarding the decision rule, we retain a classification, which consists to assign the observation at time t to the regime with the higher probability, *i.e.*:

$$i^* = ArgmaxP(S_t = i \mid y_1, \ldots, y_T).$$

5.2 Results

We now present the results of an estimated turning points chronology for the industrial production in the euro-zone by applying both parametric and non-parametric methods. We will first consider the classical cycle and then the growth cycle.

We start with the application of the three-regime Markov-switching model MS(3) on the monthly industrial production in the euro-zone (3-months growth rate).

Parameter estimates are presented in Table 2. Persistence is quite high since the probability of staying in the same regime overpass 80% in the three regimes. The transition probability from a high regime to a low regime is quasi null, which means that before entering into recession, the euro-zone industrial production first decelerates.

On the contrary, from low regime to high regime, the transition probability is 5% while it is 14% to the intermediate regime. As can be observed in Figure 10, only once, in 1975, the recovery was very strong.

The average duration of recessions is around 8 months and 5 months for high and low growth regimes. The respective unconditional probabilities (38% and 12%) are logically inversely related.

Table 2: Estimated parameters for a MSH(3)-AR(0) model applied to the IPI growth rate over 3 months from 1971-M1 to 2002-M8

	P_{i1}	P_{i2}	P_{i3}	Duration	Uncond. prob.
Regime 1	0.8139	0.1410	0.0451	5.37	0.1230
Regime 2	0.0460	0.8663	0.0877	7.48	0.4970
Regime 3	0.0001	0.1292	0.8707	7.73	0.3801

Parameters	Estim. value	S.E. estim.
μ_1	-0.0198	0.0033
μ_2	0.0018	0.0014
μ_3	0.0168	0.0016
σ_1	0.0142	
σ_2	0.0092	
σ_3	0.0100	

As concerns the intermediate regimes, the quite high unconditional probability (50%), relative to an average duration of only 7.5 months, is due to the more frequent growth cycles than business cycles over this period of time. The annualised average growth rates in low, intermediate and high regimes are -8%, 0.8% and 6.8% respectively.

The volatility is larger in the recession regime, which is a usual result. The transitory gaps in the original series (as for instance in the beginning of 1979) can be seen as outliers and entail the occurrence of regimes with a very short duration.

Since we did not correct previously for outliers, we imposed that a regime should last at least two months in order to be recognised as a cycle phase.

5.2.1 The industrial business cycle

Regarding the non parametric approach, it is worth saying that, as usual, the duration of the low phases of the business cycle seems to be only a few months, which is clearly lower than the one of high phases. Therefore, the censoring rule (iii) related to the minimum duration of a phase of the cycle strongly conditions the turning points chronology. For example, Table 3 presents the dating results for different minimum duration values.

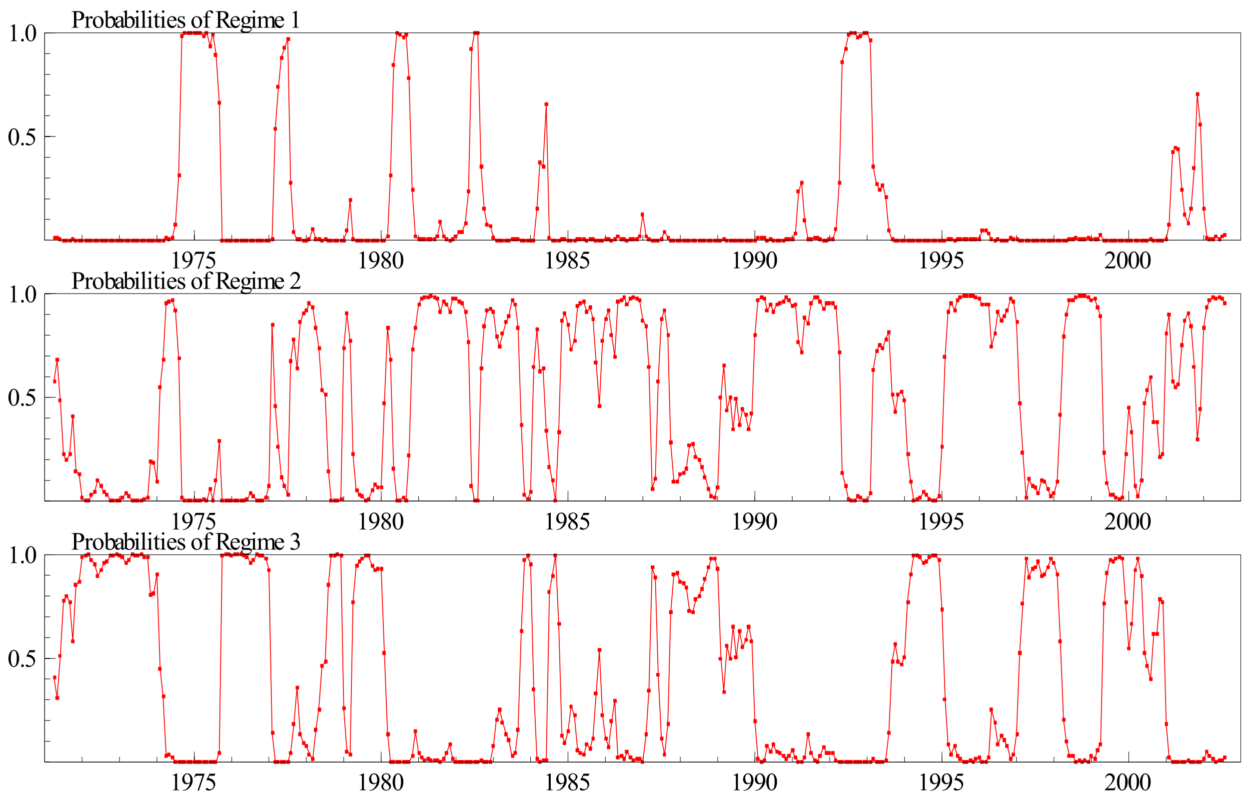

Figure 10: Smoothed probabilities of being in a given regime for monthly IPI in the euro-zone 1971-M1 to 2002-M8

Table 3: Non parametric industrial business cycle dating for the euro-zone according to the censoring rule on minimum duration of cycle phases (1971-M1 to 2002-M8)

	2 months	4 months	6 months	8 months
Peak B	M5 1974	M5 1974	M5 1974	M5 1974
Trough B	M8 1975	M8 1975	M8 1975	M8 1975
Peak B	M1 1977	M1 1977	M1 1977	
Trough C	M7 1977	M7 1977	M7 1977	
Peak B	M2 1980	M2 1980	M2 1980	M2 1980
Trough C	M10 1980	M10 1980	M10 1980	M10 1980
Peak B	M10 1981	M10 1981	M10 1981	M10 1981
Trough C	M11 1982	M11 1982	M11 1982	M11 1982
Peak B	M1 1984	M1 1984		
Trough C	M5 1984	M5 1984		
Peak B	M1 1991			
Trough C	M4 1991			
Peak B	M2 1992	M2 1992	M2 1992	M2 1992
Trough C	M6 1993	M6 1993	M6 1993	M6 1993
Peak B	M1 1996			
Trough C	M3 1996			
Peak B	M6 1998	M6 1998	M6 1998	
Trough C	M1 1999	M1 1999	M1 1999	
Peak B	M1 2001	M1 2001	M1 2001	M1 2001

Obviously, when the minimum duration is small, a lot of mini-cycles are taken into account. It is interesting to note, that a minimum duration of 8 months implies an industrial business cycle dating similar to the global business cycle dating (see for instance Anas, 2000, Harding and Pagan, 2001 or Krolzig, 2001). That is the reason why IPI is often chosen as a monthly proxy of GDP in empirical studies.

However, this proxy must be carefully considered insofar as the part of industrial production in the whole economic production is decreasing in the euro-zone. The results provided by the parametric method are very similar to those obtained with the non- parametric (with a minimum duration phase of 6 months) method, excepted at the end of the sample (see Table 4). Indeed, the industrial recession due to the Asian crisis during the second semester of 1998 is not signalled by the Markov-switching model.

Therefore, the specification of the model should be changed if we want to identify the Asian crisis. Several modifications have been used regarding the autoregressive degree, but it has not changed the dating.

However, we noticed an instability of the model. Indeed, if we limit the period of estimation to 1980-2002, then the model is able to identify the industrial recession due to the Asian crisis. Therefore, we illustrate again the idea that the dating is not a natural outcome of the Markov-switching model and there is a need to validate the model with an outside dating.

Table 4: Parametric industrial business cycle dating for the euro-zone (1971-M1 to 2002-M8)

Peak B	M8 1974
Trough C	M8 1975
Peak B	M2 1977
Trough C	M7 1977
Peak B	M4 1980
Trough C	M10 1980
Peak B	M5 1982
Trough C	M8 1982
Peak B	M4 1992
Trough C	M2 1993
Peak B	M10 2001
Trough C	M12 2001

Regarding the last recession, there is a clear peak given for January 2001 with the non-parametric method, whatever minimum duration value we select. In November 2002, there was not yet any sign of exit of the industrial recession. By contrast, the Markov-switching model has some difficulty to recognise a recession: the signal is given lately in December 2001 and did not last more than two months. A recession signal was coming at the start of 2001 but did not materialise because there was a pick-up in growth starting at spring.

Without the terrorist attack of September 2001, the industrial recession in the euro-zone might have been avoided. That external shock explains the renewed surge in the recession probability at the end of 2001, as can be seen in Figure 10.

It must also be noted that the timing of turning points is not the same in both methods. The dates of peaks in the business cycle provided by the MS model are lagged with a lag varying between 1 and 9 months, while the dates of troughs are slightly advanced.

5.2.2 The industrial growth cycle

The growth cycles extracted by using the BK and HP filters are presented in Figure 11 and are quite similar. The BK filter provides more or less the same turning points dating than the HP filter, as it can be seen in Table 5, but identifies two supplementary "mini-cycles" in 1981 and 1991.

The non-parametric dating of the growth cycle is also very close to the parametric dating obtained with the 3-regimes Markov-switching model, excepted during the 1980's when the European cycle was more erratic.

Generally, it appears that parametric and non-parametric methods tend to provide similar results when applied to growth cycle series. This is perhaps due to the fact that the growth cycle possesses a quite symmetrical behaviour, contrary to the business cycle.

Table 5: Industrial growth cycle dating for the euro-zone over the period 1971-M1 to 2002-M8

	Non parametric BK	Non parametric HP	Parametric MS model
Trough D			M5 1971
Peak A	M2 1974	M2 1974	M1 1974
Trough D	M7 1975	M8 1975	M9 1975
Peak A	M11 1976	M12 1976	M1 1977
Trough D	M12 1977	M2 1978	M7 1978
Peak A			M12 1978
Trough D			M3 1979
Peak A	M1 1980	M2 1980	M2 1980
Trough D	M3 1981		
Peak A	M10 1981		
Trough D	M12 1982	M11 1982	M9 1983
Peak A			M1 1984
Trough D			M5 1984
Peak A	M10 1985	M3 1986	M10 1984
Trough D	M5 1987	M2 1987	M9 1987
Peak A	M7 1990	M1 1991	M12 1989
Trough D	M4 1991		
Peak A	M12 1991		
Trough D	M7 1993	M7 1993	M12 1993
Peak A	M3 1995	M1 1995	M1 1995
Trough D	M12 1996	M12 1996	M1 1997
Peak A	M2 1998	M2 1998	M3 1998
Trough D	M3 1999	M3 1999	M4 1999
Peak A	M10 2000	M11 2000	M12 2000

Lastly, note that, regarding the last growth cycle, there is quite a similitude in the peak dating in the last quarter of 2000 but it seems to be too soon to date the last trough.

6 Conclusion

There is a need to define clearly the concepts of cycles, turning points and turning point detection. There is a distinction to be made between ex-post dating and detecting (the latter including real time detecting and predicting). It seems recommendable to concentrate on the follow-up of growth cycle and business cycle turning points (ABCD approach).

The traditional non-parametric approach has been challenged in the last 10 years by the new class of Markov-switching regime models.

It seems advisable for dating purpose to have an expert analysis based mainly on non-parametric methods like the Bry and Boschan (1971) algorithm, at least for the business cycle. This is due to the necessary calibration of models on dating.

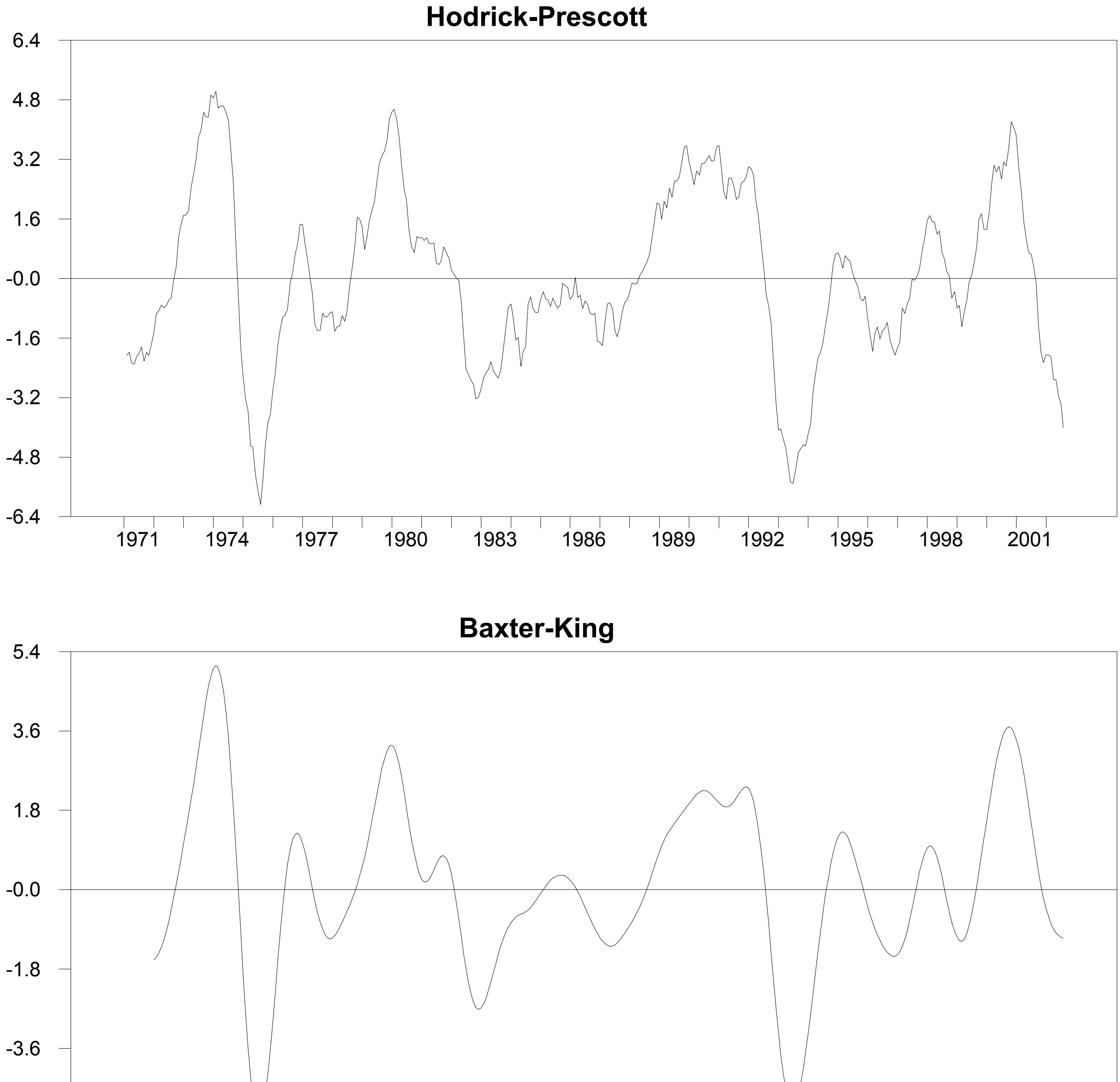

Figure 11: Growth cycle of the monthly IPI in the euro-zone from 1971-M1 to 2002-M8

However, for dating growth cycle, MS models seem more competitive. Concerning detection, those new models may be very competitive and should benefit from the new comovement extraction tools, on small or big data sets. Those models will have the purpose to detect in the most efficient way the turning points without necessarily replicating other stylized facts. From a practical point of view, we propose the recommendations synthesised in Table 6.

Table 6: Final recommendations

	Dating	Detecting
Non-parametric	Yes (for business and growth cycle)	Maybe
Parametric	No (for business cycle) and maybe for growth cycle	Yes (for business and growth cycle)

References

Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, L. Reichlin, and G. Veronese (2001): "EuroCOIN: A real time coincident indicator of the Euro area business cycle," CEPR Discussion Paper No. 3108.

Anas J. (1997): "Forecasting turning points of the French growth cycle," 22nd Conference Ciret, Helsinki, Finland.

Anas J. (2000): "Le cycle économique européen: datation et détection," in Les Mutations de l'Economie Mondiale, C. de Boissieu (ed.), Economica, Paris.

Anas J., and M. Nguiffo-Boyom (2001): "A new indicator based on Neftçi's approach for predicting turning points of the eurozone cycle," *Quarterly Journal of Economic Research*, to appear.

Anas, J., and L. Ferrara (2002a): "Un indicateur d'entrée et sortie de récession :application aux Etats-Unis," Document de travail No. 58, Centre d'Observation Economique (www.coe.ccip.fr).

Anas, J., and L. Ferrara (2002b): "Detecting cyclical turning points: the ABCD approach and two probabilistic indicators," paper presented at the 26th Ciret Conference in Taiwan, October 2002 (www.coe.ccip.fr).

Artis, M. (2002): "Dating the business cycle in Britain," Discussion paper series No. 17, CGBCR, University of Manchester.

Astolfi, R., D. Ladiray, and G. L. Mazzi (2001): "Business cycle extraction of eurozone GDP: direct versus indirect approach," *Quarterly Journal of Economic Research*, 3, pp. 377-398.

Boldin, M. D. (1996): "A check on the robustness of Hamilton's Markov Switching model approach to the economic analysis of the business cycle," *Studies in Nonlinear Dynamics and Econometrics*, 1, pp. 35-46.

Bry, G., and C. Boschan (1971): "Cyclical analysis of time series: Selected procedures and computer programs," NBER, Technical Paper, n. 20.

Burns, A. F., and W. C. Mitchell(1946): *Measuring Business Cycles*, NBER, Columbia University Press.

Chen, M. (2002): "The use of the composite leading indicators for forecasting business cycle turning points," paper presented at the 26th Ciret Conference in Taiwan, October 2002.

Clements, M. P., and H. M. Krolzig (1998): "A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP," *Econometrics Journal*, 1, pp. 47-75.

Clements, M. P., and H. M. Krolzig (2000): "Modelling business cycle features using switching regimes models," Mimeo.

Diebold, F. X., and G. D. Rudebusch (1996): "Measuring business cycles: a modern perspective," *The Review of Economics and Statistics*, 78, pp. 67-77, reprinted in: *Business Cycles: Durations, Dynamics and Forecasting*, Princeton University Press, Princeton, 1999.

Doornik, J. A. (1999): *Object-Oriented Matrix Programming using Ox*, Third Edition, Timberlake Consultant Press, London.

Franses, P. H., and D. van Dijk (2001): *Non-linear Time Series Models in Empirical Finance*, Cambridge University Press.

Hamilton J. D. (1989): "A new approach to the economic analysis of non-stationary time series and the business cycle," *Econometrica*, 57, pp. 357-384.

Hamilton J. D. (1990): "Analysis of time series subject to changes in regime," *Journal of Econometrics*, vol. 45, pp. 39-70.

Hamilton J. D. (1994): *Time Series Analysis*, Princeton University Press, Princeton.

Hamilton J. D. (1996): "Specification testing in Markov-switching time series models," *Journal of Econometrics*, 70, pp. 127-157.

Harding, D., and A. Pagan (2001a): "A comparison of two business cycle dating methods," manuscript, University of Melbourne.

Hodrick, R. J., and E. C. Prescott (1997): "Post-war US business cycles: An empirical investigation," *Journal of Money, Credit, and Banking*, 29, pp. 1-16.

Keilis-Borok, V., J. Stock, A. Soloviev, and P. Mikhalev (2000): "Pre-recession pattern of six economic indicators in the USA," *Journal of Forecasting*, 19, pp. 65-80.

Kim, C. J., and C. R. Nelson (1998): "Business cycles turning points, a new coincident index and tests of duration dependence based on a dynamic factor model with regime switching," *The Review of Economics and Statistics*, 80, pp. 188-201.

Krolzig, H. M. (1997): Markov-switching vector autoregressions. Modelling, statistical inference and applications to business cycle analysis, Springer, Berlin.

Krolzig, H. M. (2001): "Markov-Switching procedures for dating the euro-zone business cycle," *Quarterly Journal of Economic Research*, 3, pp. 339-351.

Krolzig, H. M., and J. Toro (1999): "A new approach to the analysis of shocks and the cycle in a model of output and employment," European University Institute, Working paper ECO, n99-30.

Krolzig, H. M., and J. Toro (2001): "Classical and modern business cycle measurement: The European case," Discussion Paper in Economics 60, University of Oxford.

Lahiri, K., and J. G. Wang (1994): "Predicting cyclical turning points with leading index in a Markov switching model," *Journal of Forecasting*, 13, pp. 245-263.

Layton, A. P., and D. Smith (2000): "Further evidence on the three phases of the US business cycle," *Applied Economics*, 32, pp. 1133-1143.

Lommatzsch, K., and S. Stephan (2001): "Seasonal adjustment method and the determination of turning points of the EMU business cycle," *Quarterly Journal of Economic Research*, 3, pp. 399-415.

Neftçi, S. (1982): "Optimal predictions of cyclical downturns," *Journal of Economic Dynamics and Control*, 4, pp. 307-327.

Oecd (2002): An update of the OECD composite leading indicators, Statistics Directorate, OECD, December 2002.

Okun, A. M. (1960): "On the Appraisal of Cyclical Turning-Point Predictors," *Journal of Business of the University of Chicago*, 33.

Oller, L. E., and L. Koskinen (2002): "A classifying procedure for signaling turning points," paper presented at the 26th Ciret Conference in Taiwan, October 2002.

Pagan, A. (2001): "Learning about models and their fit to data," mimeo, Australian National University.

Sichel, D. (1994): "Inventories and the three phases of the business cycle," *Journal of Business and Economic Statistics*, 12, pp. 269-277.

Stock, J. H., and M. W. Watson (1989): "New indexes of coincident and leading indicators," in: *Macroeconomics Annual*, vol. 4, MIT Press.

Watson, M. (2000): "Macroeconomic forecasting using many predictors," paper presented at the World Congress of the Econometric Society, August 2000.

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Turning Points Detection"

NON-PARAMETRIC TURNING POINT DETECTION, DATING RULES AND THE CONSTRUCTION OF THE EURO-ZONE CHRONOLOGY [1]

Don Harding[2]

The University of Melbourne

Turning points mark the onset of decisive macroeconomic change. As such they are inherently non-parametric concepts. Three markers of what might be considered as decisive change are suggested and translated into non-parametric procedures for locating turning points. These lead to the classical cycle where turning points are marked by sustained falls or rises in the level of economic activity; the growth cycle where turning points are marked by sustained falls or rises in the growth rate of economic activity about its trend rate; and the acceleration cycle where turning points are marked by sustained accelerations or decelerations in the growth rate of economic activity. There are a number of pitfalls in calculating the growth cycle as a decision the must be made about which part of the permanent component of economic activity is to be labeled as the trend. These pitfalls are identified, a rationale is given for only considering the deterministic part of the permanent component as trend when studying business cycles and diagnostic procedures are suggested to aid those who ignore this advice and seek to also treat as "trend" the stochastic part of the permanent component. Parametric methods of detecting turning points are briefly considered and procedures suggested for comparing them with those obtained from non parametric procedures. Methods for aggregating turning points are suggested. Finally, the procedures discussed in this paper are applied to develop a Euro Area chronology and the result is compared with the chronology obtained from EuroCoin.

[1]The empirical estimates in this paper are preliminary and may change if more suitable data is obtained.

[2]This paper draws on work undertaken jointly with Adrian Pagan. However, the views expressed here are my own and I am solely responsible for any errors.

1 Introduction

> **Turning point**, a stationary point on a curve that is either a maximum or a minimum. *The New Penguin Dictionary of Science*
>
> **Turning point**, point in place, time, development, etc., at which decisive change occurs. *The Concise Oxford Dictionary.*

The two definitions given above capture the essence of turning points. The *New Penguin* definition relates turning points to local maxima and minima in some ordered series Y_t. The *Concise Oxford* definition adds to this by suggesting that not all local maxima and minima should be considered as turning points rather only those at which decisive change occurs should be given that title.

What constitutes decisive change is then the key issue to be resolved prior to formalizing and implementing a definition of turning points. Three markers that are based on successively weaker notions of what constitutes decisive change are presented in Section 2 where they are also formalized into non parametric procedures for locating turning points. Many business cycle researchers become involved in unnecessary and often counterproductive efforts to remove a permanent component from the data prior to business cycle analysis. Thus Section 2 also provides a brief discussion of the pitfalls in attempts to removing the permanent component of a series prior to locating turning points in that series.

These procedures for locating turning points are also applied in Section 2 to construct GDP-based business cycle chronologies for several European economies and the Euro area as a whole. In parametric models it is common to equate decisive change with regime change. A popular example of this approach is provided by Hamilton (1989) in which a regime change occurs via a shift from a high to low trend rate of growth, the switch being turned on and off by a binary Markov process. The Markov switching (MS) approach is discussed in Section 3 where it is shown that it can also be interpreted in terms of local maxima and minima in some series. Thus, the difference between the MS and non parametric approaches lies firstly in the series in which local maxima and minima are located and secondly in the criteria that are used to determine what constitutes decisive change in that series.

Once turning points have been located in several series it is of interest to investigate whether there exists a common cycle and to study the synchronization between the common cycle and the specific cycles from which it was constructed. Section 4 sets out the procedure for aggregating turning points developed in Harding and Pagan (2000). This procedure is applied to European data in Section 5 to obtain an Euro Area reference cycle. The common cycle located by aggregating turning points is compared with the common cycle located in the aggregate of Euro area GDP. Conclusions are presented in Section 7.

2 Three non-parametric definitions of turning points

It is useful to start by considering the non-parametric description of the **specific cycle** for a single series Y_t, $t = 1..T$. If $Y_t > 0$ for all t then, it is convenient to work with $y_t = ln(Y_t)$. I will assume for convenience of exposition that $E\Delta y_t = \alpha_1$ a constant over $t = 1, .., T$ and will refer to $\tau_t \equiv y_0 + \alpha_1 t$ as the linear deterministic part of the permanent component of y_t. With these preliminaries in place I proceed in Sections 2.1, 2.2 and 2.3 to discuss the classical, growth and acceleration cycles respectively that are based on successively weaker notions of what constitutes decisive change.

2.1 The classical cycle

The strongest notion of decisive change occurs when a series changes direction so that y_t is a local maxima or minima. The dates at which these local maxima and minima occur may be represented by a pair of binary series CCP_t and CCT_t where[3]

$$CCP_t = \mathbf{1}\left(y_t > Max\{y_{t-k}, ., y_{t-1}, y_{t+1}, ..y_{t+k}\}\right)$$

$$CCT_t = \mathbf{1}\left(y_t < Min\{y_{t-k}, ., y_{t-1}, y_{t+1}, ..y_{t+k}\}\right)$$

It is useful to note that this definition can also be expressed in terms of long differences viz,

$$CCP_t = \mathbf{1}\left(0 > Max\{y_{t-k} - y_t, ., y_{t-1} - y_t, y_{t+1} - y_t, ..y_{t+k} - y_t\}\right)$$

$$CCT_t = \mathbf{1}\left(0 < Min\{y_{t-k} - y_t, ., y_{t-1} - y_t, y_{t+1} - y_t, ..y_{t+k} - y_t\}\right)$$

Turning points defined in this way are said to be from the **classical cycle** that was studied by Burns and Mitchell. The last date before y_t shows a sustained decline (rise) is referred to as the **classical cycle peak (trough)**.

This definition is sufficiently general as to accommodate business cycles arising from highly non-linear deterministic data generating processes as well as the more familiar stochastic processes in which shocks are the driving force. Where shocks are taken to be the driving force of fluctuations the **classical cycle peak** marks the onset of negative shocks that are so large relative to the trend rate of growth that the series falls for a substantial period of time. The date at which this sequence of large negative shocks ends is referred to as the **classical cycle trough**[4].

Classical cycle chronologies are maintained by:

- The National Bureau of Economic Research (NBER) for the United States.
- The Conference Board (TCB) for the United States and eight other countries, four of which are European.
- The Economic Cycle Research Institute (ECRI) for the United States and seventeen other countries seven of which are European.

[3] I use the notation that $\mathbf{1}(S)$ is an indicator statement that takes the value 1 if the statement S is true and zero otherwise.

[4] In the case of a non-linear deterministic data generating process the criteria for determining classical cycle peak and trough mark out particular regions of R^{2k+1}.

All of the classical cycle chronologies mentioned above are obtained by procedures that aggregate the turning points located in several series to obtain a reference cycle.

An algorithm to construct a reference cycle by aggregating turning points is presented in Section 4 and applied in Section 5. Classical cycle chronologies based on turning points in GDP were constructed for European Countries and the Euro Area by Harding and Pagan (2001) and are presented below.

2.1.1 Representing the cycle by a binary series

An important innovation of Wesley Mitchell's was to represent the classical cycle by a binary variable S_t that takes the value one in expansions and zero in contractions. S_t can be obtained from the binary series that locate peaks and troughs via the following recursion

$$S_t = S_{t-1}\left(1 - CCP_{t-1}\right) + \left(1 - S_{t-1}\right)_{t-1} CCT_{t-1} \tag{1}$$

where the initial value is chosen so that $S_1 = 1$ if the first turning point is a peak and $S_1 = 0$ otherwise.

The recursion (1) provides a useful tool for understanding the properties of the classical cycle.

For example, taking the unconditional expectation of S_t and making use of the notation that $E\left\{CCP_{t-1}|St-1=1\right\} = p$ and $E\left\{CCT_{t-1}|S_{t-1}=0\right\} = q$ and the assumption that $ES_t = ES_{t-1}$, yields

$$\begin{aligned}
ES_t &= ES_{t-1}\left(1 - CCP_{t-1}\right) + E\left(1 - S_{t-1}\right)_{t-1} CCT_{t-1} \\
&= Pr\left(S_{t-1}=1\right) E\left\{\left(1 - CCP_{t-1}\right)|S_{t-1}=1\right\} \\
&\quad + Pr\left(S_{t-1}=0\right) E\left\{CCT_{t-1}|S_{t-1}=0\right\} \\
&= ES_t\left(1-p\right) + \left(1 - ES_t\right) q \\
&= \frac{q}{p+q}
\end{aligned}$$

which shows how the $Pr\left(St=1\right)$ is related to the unconditional probabilities of peaks and troughs.

2.1.2 Classical cycle turning points for Euro Area GDP and selected European economies

The BBQ algorithm with $k = 2$ was used to obtain classical cycle turning points in Euro area GDP and in the GDP of selected European economies [5]. The results are summarized in Figure 1 where peaks are shown by "up-spikes" and troughs by "down spikes".

[5] Data for Euro Area GDP was taken from Fagan, Henry, and Mestre (2001). Data on individual countries was taken from the OECD data base. In the case of Germany the statistics for unified Germany prior to usinfication were constructed by splicing the West German and unified Germany series in that database. The data runs from first quarter 1970 to the last quarter 1998. This data was used because the data supplied by Eurostat was not seasonally adjusted and was therefore unsuitable for business cycle analysis.

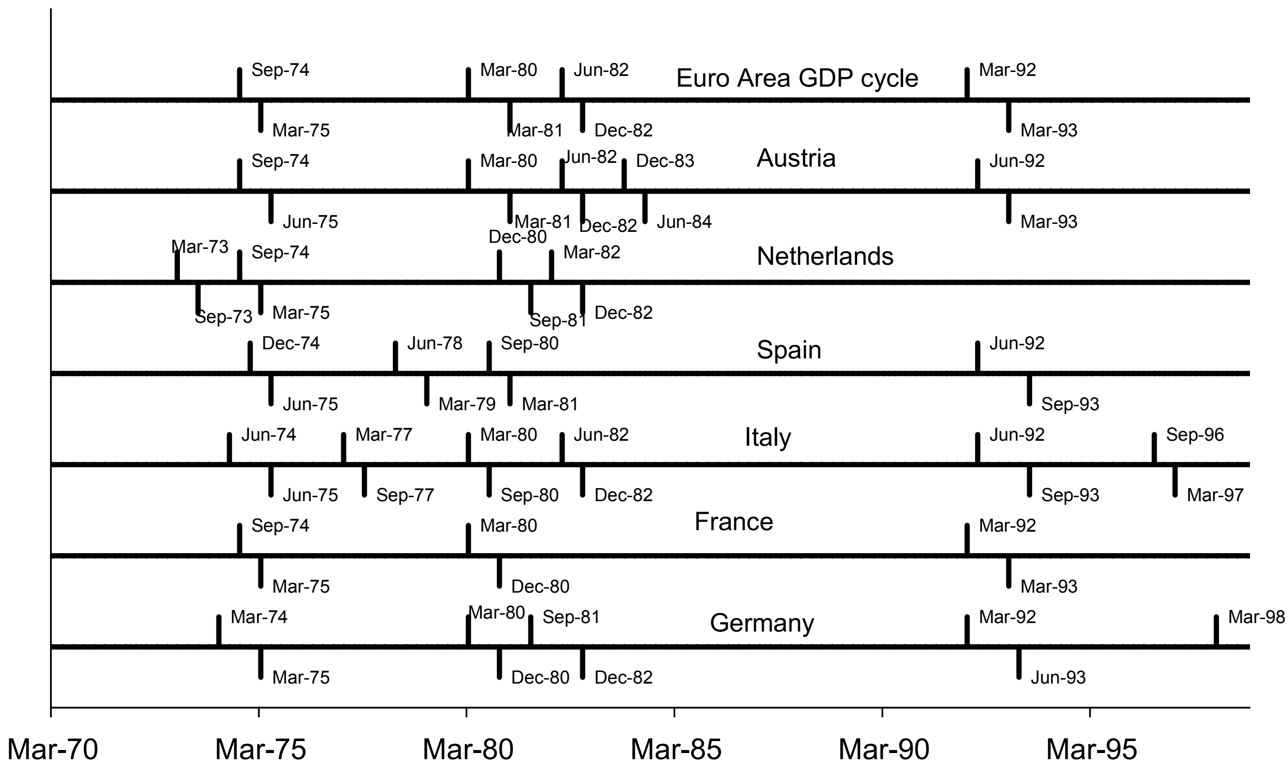

Figure 1: Classical cycle turning points in Euro Area GDP and in GDP of selected European economies, 1970.1 to 1998.4

Each turning point is labeled with the date at which it occurs. A striking feature of the data is the seeming synchronization of turning points.

The degree of synchronization can be measured by the *concordance index* that was advocated in Harding and Pagan (2002b) and that has the form

$$\hat{I} = \frac{1}{T}\{\sum_{t=1}^{T} S_{xt}S_{yt} + \sum_{t=1}^{T}(1 - S_{xt})(1 - S_{yt})\}. \tag{2}$$

Following Harding and Pagan (2000) Equation (2) can be reparameterised as

$$\begin{aligned} \hat{I} &= 1 + \frac{2}{T}\sum_{t=1}^{T} S_{xt}S_{yt} - \hat{\mu}_{S_x} - \hat{\mu}_{S_y} \\ &= 1 + 2\hat{\rho}_S\hat{\sigma}_{S_x}\hat{\sigma}_{S_y} + 2\hat{\mu}_{S_x}\hat{\mu}_{S_y} - \hat{\mu}_{S_x} - \hat{\mu}_{S_y} \end{aligned} \tag{3}$$

where $\hat{\rho}_S$ is the estimated correlation coefficient between S_{xt} and S_{yt}.

Harding and Pagan Harding and Pagan (2000) argue that since the concordance index is also monotonic in ρ_S, it is natural to shift attention away from the former to the latter i.e. to focus upon the correlation between the two states S_{xt} and S_{yt}.

They suggest that two cycles are *perfectly positively synchronized* when $\rho_S = 1$, *perfectly negatively synchronized* when $\rho_S = -1$ and *unsynchronized* when $\rho_S = 0$.

Table 1: Concordance indexes and correlations of cycles in GDP for selected European countries

	GER	FRA	ITA	SPA	NET	AU	EU
GER	$\ddots$	0.91	0.87	0.83	0.83	0.88	0.91
FRA	0.64	$\ddots$	0.90	0.91	0.87	0.94	0.97
ITA	0.52	0.52	$\ddots$	0.87	0.84	0.91	0.91
SPA	0.29	0.43	0.42	$\ddots$	0.84	0.90	0.90
NET	0.27	0.14	0.22	0.10	$\ddots$	0.88	0.90
AU	0.53	0.68	0.59	0.48	0.36	$\ddots$	0.97
EU	0.67	0.85	0.58	0.44	0.40	0.83	$\ddots$
$\widehat{\mu}_S$	0.83	0.92	0.85	0.90	0.91	0.88	0.90
$\widehat{\sigma}_S$	0.14	0.07	0.13	0.09	0.08	0.10	0.09

Hence one will generally wish to test if $\rho_S = 0$ using $\hat{\rho}_S$. They also suggest a test of synchronization for several series viz suppose there are series $x_{1t}, ..., x_{nt}$ which will have associated specific cycles $S_{jt}, j = 1, ..., n$.

Let the correlation matrix of the n series, S_{jt}, be $C_{1...n}$. Then perfect synchronization of the n cycles would have $|C_{1...n}| = 0$ and no synchronization would have $|C_{1...n}| = 1$.The statistics $\left\{\widehat{I}, \widehat{\rho}_S, \widehat{\mu}_S, \widehat{\sigma}_S\right\}$ are reported in Table 1; the concordance statistic $\widehat{I}$ is above the diagonal while $\widehat{\rho}_S$ is below the diagonal, $\widehat{\mu}_S$ and $\widehat{\sigma}_S$ are in the bottom two rows of the table. Reported values of the concordance statistic are large but caution needs to be exercised in interpreting this result since it may reflect the high probabilities of being in expansion. The correlation are immune from that criticism and are somewhat reduced in magnitude. But caution needs to also be used in interpreting the $\widehat{\rho}_S$ since S_{xt} and S_{yt} are serially correlated and heteroscedastic thus allowance must be made for these features of the data generating process when constructing tests about the degree of synchronization.

Procedures for construction heteroscedasticity and autocorrelation (HACC) robust estimates of the variance of $\widehat{\rho}_S$ are discussed in Harding and Pagan (2000). Standard and HACC robust t-statistics for the null of no bivariate synchronization are reported in Table 2 where the HACC tests use a Bartlet window and a lag length of 10 quarters. The standard t-statistics are above the diagonal and the HACC robust t-statistics are below the diagonal. Clearly, the correction for serial correlation makes a great difference; the robust t-statistics suggest that there is weak evidence for bivariate synchronization in the group of four countries Germany, France, Italy and Austria. The determinant of the correlation matrix for the six countries is 0.11 with p-value=0.14 which means that we cannot reject the null hypothesis of no synchronization for the Euro Area classical cycle as a whole.[6]. The determinant of the sub correlation matrix for Germany, France, Italy and Austria is 0.18 with p-value=0.07. This later statistic suggests that there is weak evidence for a common classical cycle for these four countries. The existence and nature of this common cycle will be explored further in Section 5 where these turning points are aggregated to obtain a Euro Area classical reference cycle.

[6]The p-value is based on the simulated distribution of the determinant. See Harding and Pagan (2000) for details.

Table 2: Standard and robust t-statistics for the null hypothesis of non correlation of classical cycle states in GDP for selected countries

	GER	FRA	ITA	SPA	NET	AU	EU
GER	$\ddots$	8.86	6.57	3.32	2.98	6.76	9.70
FRA	1.90	$\ddots$	6.52	5.14	1.53	10.10	17.66
ITA	2.10	1.90	$\ddots$	4.98	2.43	7.97	7.66
SPA	1.39	1.66	1.32	$\ddots$	1.05	5.9	5.31
NET	1.39	0.70	1.31	1.04	$\ddots$	4.12	4.70
AU	2.13	1.89	2.12	1.83	1.68	$\ddots$	16.03
EU	2.13	1.88	2.14	1.79	1.63	2.03	$\ddots$

2.2 The growth cycle

A weaker notion of what constitutes decisive change occurs where the growth rate of the series moves above or below the trend rate of growth α_1. The turning points of the growth cycle are defined by the pair of binary series GCP_t and GCT_t

$$GCP_t = \mathbf{1}\left(0 > Max\{\frac{y_{t-i} - y_t}{i} - \alpha_1, ..., \frac{y_{t+i} - y_t}{i} + \alpha_1, \text{ for } i = 1, .., k\}\right) \tag{4}$$

$$GCT_t = \mathbf{1}\left(0 < Min\{\frac{y_{t-i} - y_t}{i} - \alpha_1, ...\frac{y_{t+ki} - y_t}{i} + \alpha_1, \text{ for } i = 1, .., k\}\right) \tag{5}$$

The growth cycle is so named because growth cycle peaks (troughs) chronicle the dates at which growth moves from above (below) the trend rate to below (above) the trend rate of growth. Since $z_t = y_t - T_t$ equations 4 and 5 can be rewritten in terms of local maxima and minima in the deviation from trend series z_t, viz

$$GCP_t = \mathbf{1}\left(z_t > Max\{z_{t-k}, ., z_{t-1}, z_{t+1}, ..z_{t+k}\}\right) \tag{6}$$

$$GCT_t = \mathbf{1}\left(z_t < Min\{z_{t-k}, ., z_{t-1}, z_{t+1}, ..z_{t+k}\}\right) \tag{7}$$

The growth cycle was first studied by Ilse Mintz Mintz (1969), Mintz (1972) and Mintz (1974). Interpreted in terms of a stochastic data generating process **growth cycle** turning points chronicle dates at which there is an onset and then cessation of negative shocks that are so large relative to zero that the detrended series falls for a substantial period of time implying that the growth rate of the series also falls below the trend rate of growth. The last date before the onset of these negative shocks is the **growth cycle peak** and the last date before they cease is the **growth cycle trough**.

The Centre for International Business Cycle Research (CIBCR) and is successor ECRI used to maintain a growth cycle chronology and at various times the NBER has maintained a growth cycle chronology. Few chronologies of the growth cycle are now maintained on a regular basis and neither ECRI nor the NBER now maintain such a chronology. The main difficulty in maintaining a growth cycle chronology relates to the selection of the procedures to be used for estimating the trend T_t that is to be removed from y_t to yield the deviation from trend z_t. ECRI and CIBCR used the phase average trend (PAT) method that was developed at the NBER by Boschan and Ebanks (1978). The PAT procedure computes a rough trend by interpolating a straight line between the midpoints of classical cycle phases and then "smooths" out the kinks in the rough trend via a centred moving average.[7] This method of construction means that computation of the phase average trend at t can only be computed after the next classical turning point after t has been located a feature that is limiting for timely calculation of growth cycle turning points

There are several feasible alternatives to the phase average trend method. The method I prefer is as follows, assume that $E\Delta y_t$ is constant for $t = 1, ..., T$ and employ a linear time trend with growth rate $E\Delta y_t$. In some cases it is desirable to control for time variation in rates of population growth in this case if N_t is the population and $n_t = ln(N_t)$ then, assuming that $E_0(\Delta y_t - \Delta n_t)$ is a constant it may be useful to use $T_t = n_t + E_0(\Delta y - \Delta n)\, t$ as a measure of trend. In this vein one could also seek to adjust the trend for changes in productivity and or changes in the participation rate. The key points here are firstly that in most cases it is neither necessary nor desirable to go beyond removal of the deterministic part of the permanent component when undertaking business cycle analysis and secondly where a stochastic part of the permanent component is removed it is helpful if that component is easily identified with some economically meaningful concept such as population or labour productivity.

Attempts to do more than that advocated above to extract the permanent component from time series is an unnecessary diversion and a source of confusion in business cycle analysis. However, such attempts at removing the permanent component are commonplace and the next sub section therefore provides a digression on the pitfalls in such procedures and suggests some diagnostics to avoid those pitfalls.

2.2.1 A digression on the pitfalls in filtering and detrending data for business cycle analysis

Typically attempts at removing a time varying trend start with a decomposition of y_t into $T_t + C_t + I_t$ where these components are "trend", "cycle" and "irregular" respectively. The "trend" is then identified with the permanent component and various filters have been suggested to remove the permanent component, see Hodrick and Prescott (1997), Christiano and Fitzgerald (1998), Corbae, Ouliaris, and Phillips (2002). There are two complaints that one might make about this practice of filtering out the permanent component. The more important complaint is that several theories, most notably that associated with the RBC school, hold as a core belief that permanent shocks play a role in causing the business cycle. Removing the permanent component means that we are no longer studying the business cycle envisaged by these models. The less significant complaint is that the permanent component is not unique since the sum of any I(0) series and an I(1) series is I(1). This lack of uniqueness results in a proliferation of filtering methods. All of this is unnecessary since one can study the classical cycle in the level

[7] Computer code in the language GAUSS is available from the author to construct a phase average trend using the method described in Boschan and Ebanks (1978). The phase average trend procedure is applied to European data in Harding and Pagan (2001) where the growth cycle obtained via several detrending procedures are compared.

of the series which does not require any filtering or one can study the growth cycle using the procedures discussed above to remove only the deterministic part of the permanent component or one can study the acceleration cycle that is defined in Section 2.3 below.

Since there is considerable emphasis placed on filtering and detrending techniques at this colloquium it is worth point to the pitfalls in that approach. A common approach is to employ the following "structural" model[8]

$$y_t = y_t^p + z_t \tag{8}$$
$$y_t^p = \mu_t + y_{t-1}^p + \epsilon_t^y \qquad , var\left(\epsilon_t^y\right) = \sigma_y^2 \tag{9}$$
$$\mu_t = \mu_{t-1} + \epsilon_t^\mu \qquad , var\left(\epsilon_t^\mu\right) = \sigma_\mu^2 \tag{10}$$
$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \epsilon_t^z \qquad , var\left(\epsilon_t^z\right) = \sigma_z^2 \tag{11}$$

where z_t is identified with the "cycle" and y_t^p with the "trend". The naming of (8) to (11) as a "structural" model suggests that some other source of information has been used to identify the trend and cycle. To study whether this interpretation is is warranted Pagan (2002) examines a particular example of this model where the estimated parameters are given in Table 3.

Pagan simulated this model and applied the BBQ algorithm to the simulated data for three series $\{y_t, y_t^p, z_t\}$. The results are reported in Table 4.

It is evident that

> ... in terms of cycle durations there is absolutely no difference between the cycles in any of the series so it seems quite misleading to label one of them z_t as "cycle" and the other y_t^p as "trend". The only distinction is in their amplitudes and that simply depends upon the relative variances of each series. Since
>
> $$var\left(y_t\right) > var\left(z_t\right) > var\left(y_t^p\right)$$
>
> so to are the amplitudes ranked. Source: Pagan (2002).

Thus, if filtering procedures are to be employed to remove the permanent component of a series prior to locating turning points then a minimum requirement should be to use the BBQ algorithm as a diagnostic tool to check whether the researcher is warranted to attach the "trend" label to the permanent component and the "cycle" label to the difference between y_t and the permanent component.

Table 3: Estimated parameters for the "structural" model

$\widehat{\sigma}_y^2$	$\widehat{\sigma}_\mu^2$	$\widehat{\sigma}_z^2$	$\widehat{\phi}_1$	$\widehat{\phi}_2$
0.607	0.00091	0.301	1.401	-0.531

[8] This example is taken from Pagan (2002).

Table 4: Simulated properties of "trend", "cycle" and actual series for "structural" model

	y_t	y_t^p	z_t
Contraction duration	5.5	5.5	5.5
Expansion duration	5.5	5.5	5.5
Contraction amplitude	-3.2	-1.9	-2.6
Expansion amplitude	3.2	1.9	2.6

2.2.2 Growth cycle turning points for Euro Area GDP and selected European economies

A linear deterministic trend was removed from Euro Area GDP and the GDP of selected European economies, the procedures discussed above were applied to obtain the growth cycle for each of these series. The results are summarized in Figure 2. The concordance indices and the correlations between growth cycle states are reported in Table 5 below using the same format of presentation as for Table 5. These suggest quite high degrees of synchronization but as discussed in the case of the classical cycle one needs to make allowance for serial dependence and heteroscedasticity.

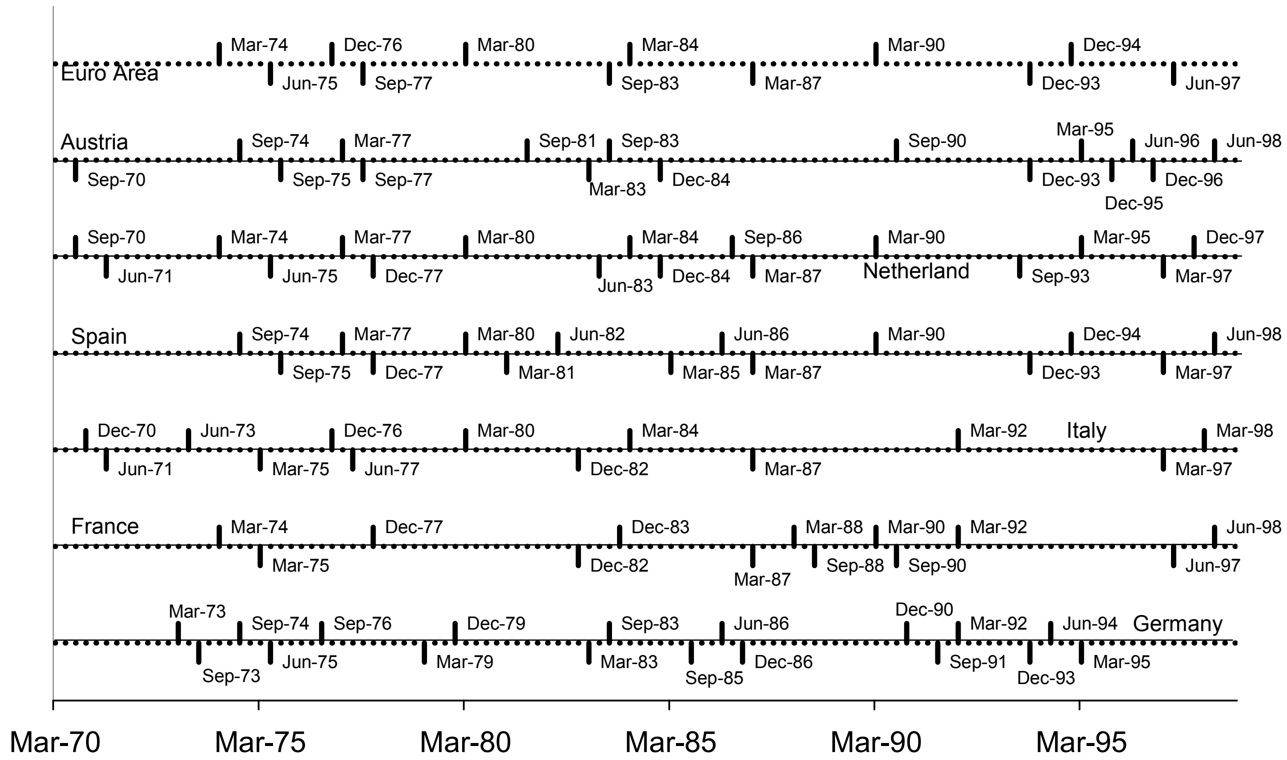

Figure 2: Growth cycle turning points for Euro Area GDP and selected European economies, 1970.1 to 1998.4

Table 5: Concordance indexes and correlations of growth cycles in GDP for selected European countries

	GER	FRA	ITA	SPA	NET	AU	EU
GER	$\ddots$	0.66	0.66	0.67	0.62	0.66	0.68
FRA	0.34	$\ddots$	0.80	0.63	0.63	0.53	0.73
ITA	0.31	0.61	$\ddots$	0.66	0.72	0.60	0.78
SPA	0.34	0.27	0.31	$\ddots$	0.83	0.80	0.83
NET	0.24	0.27	0.45	0.65	$\ddots$	0.75	0.83
AU	0.31	0.11	0.19	0.60	0.51	$\ddots$	0.70
EU	0.38	0.47	0.55	0.66	0.66	0.42	$\ddots$
$\widehat{\mu}_S$	0.56	0.45	0.51	0.56	0.53	0.66	0.49
$\widehat{\sigma}_S$	0.25	0.25	0.25	0.25	0.25	0.23	0.25

Standard and HACC robust t-statistics, constructed in line with the earlier discussion, are reported in Table 6 presented on the same basis as in Table 2. Again the HACC correction matters a lot as evidenced by the difference between the standard and HACC robust t-statistics. The evidence on the bivariate tests is somewhat more strongly in favour of synchronization of Euro Area growth cycles than it was for synchronization of classical cycles. This conclusion is supported by the determinant of the correlation matrix which is 0.135 with p-value= 0.006. Thus in later sections I will extract via non parametric methods a common growth cycle for the Euro Area as a whole.

Table 6: Standard and robust t-statistics for the null hypothesis of non correlation of growth cycle states in GDP for selected countries

	GER	FRA	ITA	SPA	NET	AU	EU
GER	$\ddots$	3.95	3.52	3.83	2.63	3.48	4.48
FRA	2.29	$\ddots$	8.26	3.08	2.97	1.16	5.68
ITA	2.00	2.80	$\ddots$	3.52	5.40	2.13	7.13
SPA	2.18	1.71	2.00	$\ddots$	9.33	8.07	9.52
NET	1.54	1.71	2.47	2.96	$\ddots$	6.38	9.39
AU	1.80	0.82	1.69	2.68	2.64	$\ddots$	5.03
EU	2.29	2.31	2.68	2.90	2.77	2.28	$\ddots$

2.3 The acceleration cycle

The weakest notion of decisive change considered in this paper occurs where the growth rates accelerates or decelerates so that letting $w_t = y_t - y_{t-l}$ be the l period growth rate of y_t, turning points are defined mathematically by the pair of binary series ACP_t and ACT_t

$$ACPt = \mathbf{1}\left(w_t > Max\{w_{t-l}, ., w_{t-1}, w_{t+1}, ..w_{t+l}\}\right)$$

$$ACTt = \mathbf{1}\left(w_t < Min\{w_{t-l}, ., w_{t-1}, w_{t+1}, ..w_{t+l}\}\right)$$

Rewriting and in terms of long differences yields

$$ACPt = \mathbf{1}\left(0 > Max\{w_{t-l} - w_t, ., w_{t-1} - w_t, w_{t+1} - w_t, ..w_{t+l} - w_t\}\right)$$

$$ACTt = \mathbf{1}\left(0 < Min\{w_{t-l} - w_t, ., w_{t-1} - w_t, w_{t+1} - w_t, ..w_{t+l} - w_t\}\right)$$

Thus the acceleration cycle gets its name because the peak measures the date at which growth rates stop increasing and begin to decrease.

Similarly, the trough measures the date at which growth rates stop decreasing and begin to increase.[9] The acceleration cycle turning points chronicle dates that mark the onset and then cessation of negative shocks that are sufficient to cause the growth rate to accelerate The last date before (after) the onset (cessation) of these negative shocks is referred to as the **acceleration cycle peak (trough)**.

The acceleration cycle has become rather popular. The Economic Cycle Research Institute (ECRI) dates such a cycle for 18 economies including seven European economies. However, ECRI uses the name growth rate cycle rather than acceleration cycle for this chronology adopting the definition that

> "growth rate cycle downturns are pronounced, pervasive and persistent declines in the growth rate of aggregate economic activity. The procedures used to identify peaks and troughs in the growth rate cycle are entirely analogous to those used to identify business cycle turning points, except that they are applied to the growth rates of the same time series, rather than their levels"
> Source: http://www.businesscycle.com/research/intlcycledates.php#.

This terminology is undesirable as it runs the risk of creating confusion with the growth cycle as defined earlier.

An example of such confusion is provided by the authors of EuroCoin who state that

> We define a recession, for example, 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 growth. Troughs and peaks are defined as the ending points of expansions and recessions respectively (sic), ie as points of minimal or maximal growth.
> Source: http://www.cepr.org/Data/eurocoin/indepth/

Now this is clearly a definition of an acceleration cycle in my terminology or a growth rate cycle in the ECRI terminology but the authors of EuroCoin go on to say

[9] The Economic Cycle Research Institute, New York, refers to the acceleration cycle as the cycle in the growth rate. However, this terminology runs the risk of being confused with the growth cycle.

Our definition of recession, expansion, peaks and troughs correspond to the so called "growth cycle".
Source: http://www.cepr.org/Data/eurocoin/indepth/

Which is clearly in conflict with their earlier statement.

2.3.1 Acceleration cycle turning points for Euro Area GDP and selected European economies

Acceleration cycle turning points for selected European countries and for Euro Area GDP are presented in Figure 3. The concordance indices and the correlations between acceleration cycle states are reported in Table 7 on the same basis as for Table 5. Clearly, on both measures there is substantial synchronization, in the sample, between the acceleration cycles of Euro Area economies. However, as with the other cycles we need to adjust for serial dependence before coming to a strong view as to the extent of synchronization.

Standard and HACC robust t-statistics for the null hypothesis of no synchronization are reported in Table 8 where there is evidence of a common Euro Area acceleration cycle. The determinant of the correlation matrix for the six countries is 0.342 with p-value 0.013 which again provides support for a Euro area acceleration cycle. In later sections an attempt will be made to extract this common acceleration cycle.

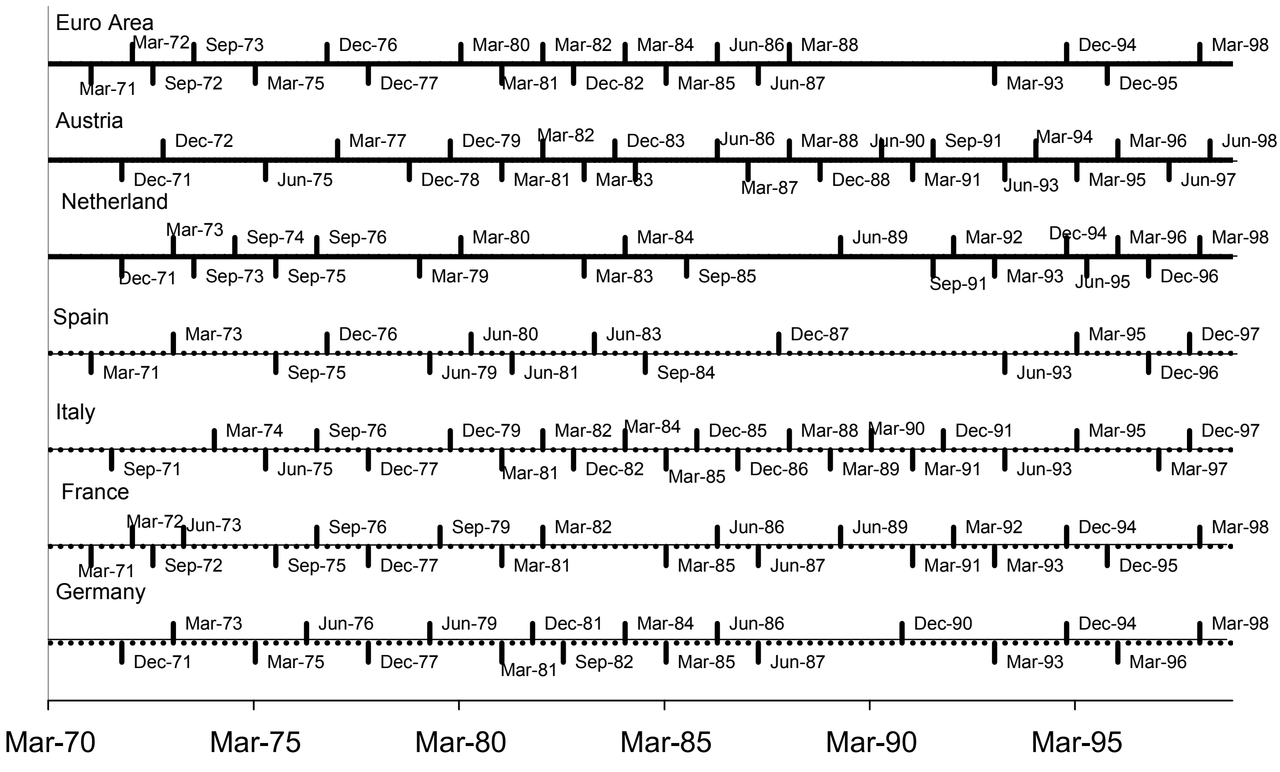

Figure 3: Acceleration cycle for Euro Area GDP and for selected European economies, 1970.1 to 1998.4

Table 7: Concordance indexes and correlations of acceleration cycles in GDP for selected European countries

	GER	FRA	ITA	SPA	NET	AU	EU
GER	$\ddots$	0.74	0.71	0.56	0.63	0.63	0.76
FRA	0.48	$\ddots$	0.70	0.63	0.66	0.61	0.82
ITA	0.41	0.39	$\ddots$	0.66	0.64	0.70	0.75
SPA	0.13	0.25	0.33	$\ddots$	0.66	0.63	0.70
NET	0.27	0.32	0.29	0.32	$\ddots$	0.63	0.63
AU	0.27	0.21	0.39	0.25	0.25	$\ddots$	0.63
EU	0.52	0.64	0.50	0.40	0.25	0.25	$\ddots$
$\widehat{\mu}_S$	0.53	0.50	0.52	0.45	0.48	0.48	0.52
$\widehat{\sigma}_S$	0.25	0.25	0.25	0.25	0.25	0.25	0.25

2.4 The choice of window width k

So far we have seen that one dimension of the concept of decisive change can be framed in terms of whether level, deviation from trend, or growth rate of a series that shows a fall. A second dimension can be framed in terms of how sustained are the movements in the series. This is captured by the window width k. The choice of k can be made on various grounds. Harding and Pagan (2002b) suggest choosing k with reference to the work of Bry and Boschan (1971) who designed a dating algorithm to formalize the procedures of the NBER business cycle dating committee. In the Bry Boschan program $k = 5$ for monthly data and thus Harding and Pagan (2002b) choose to set $k = 2$ for quarterly data in their Bry Boschan Quarterly (BBQ) program[10].

Table 8: Standard and robust t-statistics for the null hypothesis of non correlation of acceleration cycle states in GDP for selected countries

	GER	FRA	ITA	SPA	NET	AU	EU
GER	$\ddots$	5.84	4.75	1.41	2.97	2.97	6.39
FRA	2.30	$\ddots$	4.52	2.75	3.59	2.32	8.89
ITA	2.55	2.30	$\ddots$	3.67	3.18	4.55	6.10
SPA	0.88	1.81	2.12	$\ddots$	3.57	2.71	4.61
NET	2.01	1.96	2.35	2.03	$\ddots$	2.72	2.75
AU	1.83	1.54	2.38	1.78	2.18	$\ddots$	2.75
EU	2.34	2.95	2.65	2.52	1.72	1.64	$\ddots$

[10] A gauss program BBQLite that allows the user to choose the definition of a turning point, the window width k and the censoring procedures is available on request from the author.

2.5 Censoring procedures

The choice of window width k is not sufficient to ensure that only sustained movements in a series are classified as turning points. The problem can be illustrated with reference to the classical cycle trough. Suppose that $y_t < Min\{y_{t-k}, ., y_{t-1}, y_{t+1}, ..y_{t+k}\}$ and thus the formulas set out earlier seeks to locate a trough at date t. Now it is possible that $y_{t+1} > Max\{y_{t-k+1}, ., y_t, y_{t+2}, ..y_{t+k+1}\}$ and thus the definitions given earlier would suggest locating a peak at $t+1$ now if the period of the data were annual this might be entirely appropriate but if the data is quarterly or monthly then it is unlikely that anyone would consider the movement sufficiently decisive as to mark a turning point. For this reason it is necessary to apply censoring procedures to ensure that expansion and contraction phases of the business cycle have a minimum length of k periods. The BBQ program extends to quarterly data the gauss code developed by Watson (1994) for censoring business cycle chronologies in the manner of Bry and Boschan. These procedures are discussed in Harding and Pagan (2002b).

2.6 Stylized relation between classical, growth and acceleration turning points

Figure 4 provides a stylized representation of the three types of turning point identified above. As can be seen from Figure 4 each of these concepts measures a particular aspect of the business cycle.

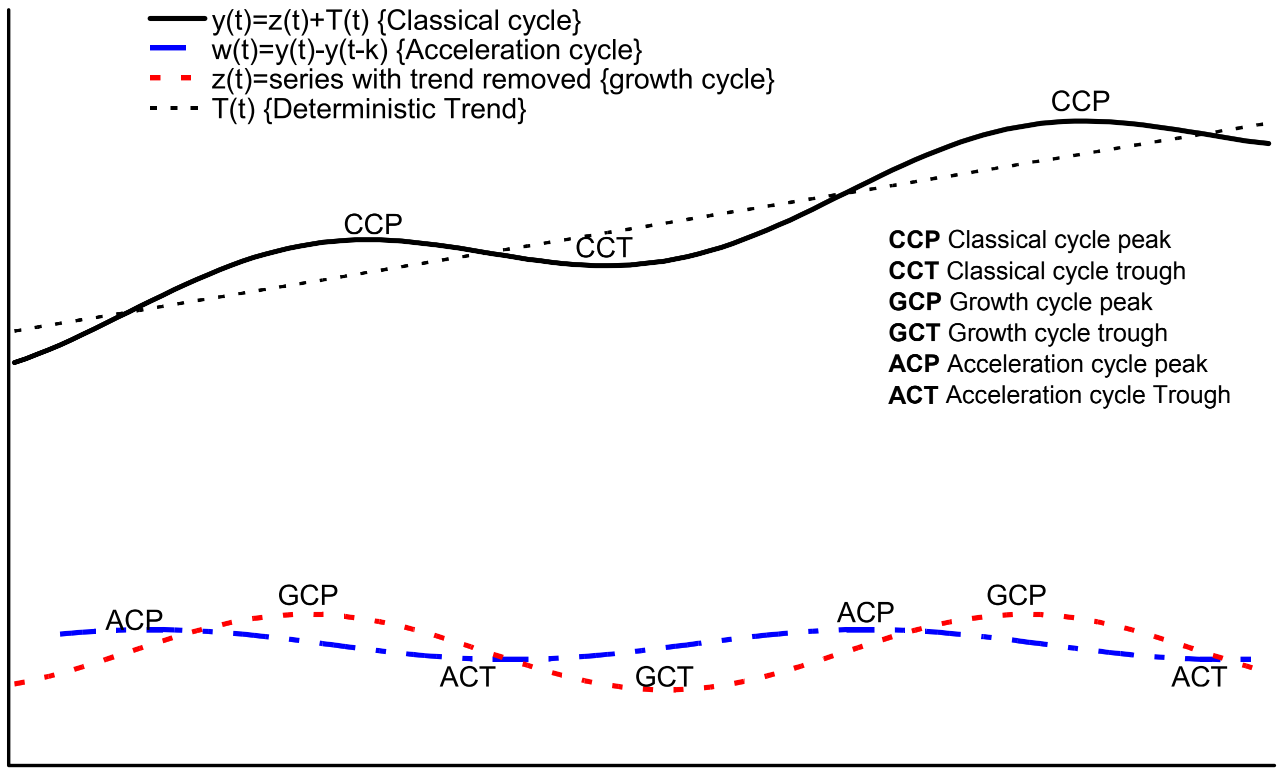

Figure 4: Stylized representation of classical, growth and acceleration cycle turning points

3 Parametric definitions of turning points

The non parametric approach to turning point detection given in the last section has been questioned by Diebold and Rudebusch who state that

> Because it is only within a regime-switching framework that the concept of a turning point has intrinsic meaning. ....In linear frameworks, by way of contrast, there are no turning points or switch times, in probabilistic structure. **One can, of course define turning points in terms of features of sample paths, but such definitions are fundamentally ad hoc**" (emphasis added) Because it is only within a regime-switching framework that the concept of a turning point has intrinsic meaning. ....In linear frameworks, by way of contrast, there are no turning points or switch times, in probabilistic structure. **One can, of course define turning points in terms of features of sample paths, but such definitions are fundamentally ad hoc**" (emphasis added) Diebold and Rudebusch (1999)

The essence of the complaint just cited is that non-parametric definitions of turning points lack generality — the dictionary meaning of ad hoc. Harding and Pagan Harding and Pagan (2002a) explore the validity of this position by examining the simplest version of a regime switching model popularized in Hamilton (1989). The model is

$$\Delta y_t = \mu_0 + \mu_1 \zeta_t + \sigma e_t \tag{12}$$

$$\Pr\left(\zeta_t = 1 | \zeta_{t-1} = 1\right) = p \tag{13}$$

$$\Pr\left(\zeta_t = 1 | \zeta_{t-1} = 0\right) = q \tag{14}$$

where $\zeta_t = 1$ denotes the high growth state $(\mu_0 < \mu_1)$ which is taken to be an expansion while $\zeta_t = 0$ denotes contraction. The parameters μ_0, μ_1, σ, p and q are estimated from the data. Filtered $Pr\left(\zeta_t = 1 | \{\Delta y_{1,..,}\Delta y_t\}\right)$ and smoothed $Pr\left(\zeta_t = 1 | \{\Delta y_t\}_{t=1}^{T}\right)$ probabilities of being in expansion are then produced as a by product of the estimation procedure. The smoothed probabilities can be used to construct turning points via a rule that a turning point occurs if $Pr\left(\zeta_t = 1 | \{\Delta y_t\}_{t=1}^{T}\right)$ crosses over some critical value which is typically chosen to be one-half. Let the pair of binary variables MSP_t and MST_t take the values 1 in peaks and troughs respectively and zero elsewhere. Then,

$$\begin{aligned} MSP_t &= \mathbf{1}\{\Pr(\zeta_{t+1} = 1 | \{\Delta y_t\}_{t=1}^{T}) < 0.5, \Pr(\zeta_t = 1 | \{\Delta y_t\}_{t=1}^{T}) > 0.5\} \\ MST_t &= \mathbf{1}\{\Pr(\zeta_{t+1} = 1 | \{\Delta y_t\}_{t=1}^{T}) > 0.5, \Pr(\zeta_t = 1 | \{\Delta y_t\}_{t=1}^{T}) < 0.5\}. \end{aligned} \tag{15}$$

3.1 The relationship between parametric and non parametric definitions of turning points

The $\{MSP_t, MST_t\}$ are the parametric analogues of the $\{CCP_t, CCT_t\}$ that were obtained via the non parametric methods described earlier. This suggests that we might seek to find simple expression for the parametric Markov switching rules so that they can be compared with the non parametric rules obtained earlier. Harding and Pagan Harding and Pagan (2002a) show how to do this. For the specific case where the parameter estimates are p= they show that a first order approximation to the MS rule () is

$$
\begin{aligned}
&peak\ at\ \ t\ if\ \left\{\begin{array}{c} 0.15\Delta y_{t+1} + 0.06(\Delta y_t + \Delta y_{t+2}) < 0.22, \\ 0.15\Delta y_t + 0.06(\Delta y_{t-1} + \Delta y_{t+1}) > 0.22 \end{array}\right\} \\
&trough\ at\ \ t\ if\ \left\{\begin{array}{c} 0.15\Delta y_{t+1} + 0.06(\Delta y_t + \Delta y_{t+2}) > 0.22, \\ 0.15\Delta y_t + 0.06(\Delta y_{t-1} + \Delta y_{t+1}) < 0.22 \end{array}\right\}
\end{aligned}
\tag{16}
$$

Evidently, to a first order of approximation the Hamilton MS dating rule is based on growth rates of GDP and the window width is 2 quarters — the same as that used in the non parametric BBQ rule that was discussed earlier. While there is no simple relation between the BBQ rule and the first order approximation to the MS rule the basis for a comparison can be made by writing $0.15\Delta y_{t+1} + 0.06(\Delta y_t + \Delta y_{t+2}) < 0.22$ as

$$
0.03\Delta y_{t+1} + 0.06\left(\Delta y_{t+1} + \Delta_2 y_{t+2}\right) + 0.06\Delta y_t < 0.22 \tag{17}
$$

The sufficient condition for the BBQ rule to locate a peak at t is that $\Delta y_{t+1} < 0$ and $\Delta y_{t+1} + \Delta_2 y_{t+2} < 0$ thus, a sufficient condition for the two rules to give exactly the same turning point would be $0.03\Delta y_{t+1}+\ 0.06\Delta y_t < 0.22$. In many instances this condition will be met but one can imagine circumstances where Δy_{t+1} is negative but small while Δy_t is large and positive and thus the two rules will produce different outcomes in these circumstances.

Interest also centres on the quality of the approximation to the MS dating rule. It is shown in Harding and Pagan (2002a) that for the parameters obtained for US GDP a second order approximation will be an accurate estimate of the smoothed the probability that the economy is in expansion. That second order approximation yields the rule that one should place a peak at date t if (the rule for a trough is symmetric)

$$
0.16\Delta y_{t+1} + 0.07\left(\Delta y_t + \Delta y_{t+2}\right) + 0.02\left(\Delta y_{t-1} + \Delta y_{t+3}\right) < 11
$$

and

$$
0.16\Delta y_t + 0.07\left(\Delta y_{t-1} + \Delta y_{t+1}\right) + 0.02\left(\Delta y_{t-2} + \Delta y_{t+2}\right) > 11
$$

Notice that the second order approximation to the MS rule includes Δy_{t+3} and thus involves a wider window width than does the BBQ rule or the first order approximation to the MS rule. Moreover, higher order approximations will involve larger window widths. The important practical implication of this is that as new data points are added to the sample $\Pr(\zeta_{T-\delta} = 1|\{\Delta y_t\}_{t=1}^{T+1})$ may well "wriggle"about thus those who seek to date the cycle in real time with MS techniques will typically need to wait longer before calling a turning point than will agencies that use non parametric methods such as the BBQ rule.

3.2 An illustrative example using the MS model to date the sunspot cycle

Some further indication of the methodological and practical problems with the using MS dating methods can be considered using the MS model to date the sunspot cycle. The cycle in the average number of sunspots is well know and is shown in Figure 5. It is evident that the graph of the sunspot number shows a number of decisive changes of direction that we would consider as turning points It is therefore of some interest to explore how adequately the parametric MS model detects these turning points.

To do this the following MS model was fit to the sunspot data yielding the parameters in Table 9.

$$
\begin{aligned}
y_t - \mu_t &= \phi_1 \left(y_{t-1} - \mu_{t-1}\right) + \phi_1 \left(y_{t-2} - \mu_{t-2}\right) + \sigma_t e_t \qquad e_t \tilde{} N(0,1) \\
\mu_t &= \mu_0 (1 - \zeta_t) + \mu_1 \zeta_t \\
\sigma_t &= \sigma_0 (1 - \zeta_t) + \sigma_1 \zeta_t \\
p_{11} &= \Pr\left(\zeta_t = 1 | \zeta_{t-1} = 1\right) \\
p_{00} &= \Pr\left(\zeta_t = 0 | \zeta_{t-1} = 0\right)
\end{aligned}
$$

The smoothed probability of being in the high growth state is plotted together with the sunspot number in Figure 6. It is evident that the MS approach does reasonably well at detecting turning points in the sunspot number for the 18^{th} century but does rather poorly in the 19^{th} and 20^{th} century. This serves to provide a warning that even though the MS approach my prove adequate in sample at detecting turning points there is no warrant that it will prove to be adequate in the future.

In summary, the MS approach suffers from several defects. First, it is not very transparent although the procedures that are described briefly above and more fully in Harding and Pagan (2002a) serve to make the MS approach more transparent in the sense that they allow it to be compared to the more transparent non parametric approach discussed in Section 2 above. Second, when the approximate parametric rules are discovered they are not very intuitive unlike the non parametric ones. Third, the MS rules are very model and sample specific features that are undesirable in producing a business cycle chronology.

Table 9: Estimated parameters for markov switching model of sunspot number

Parameter	μ_0	μ_1	ϕ_1	ϕ_1	σ_0	σ_1	p_{00}	p_{11}
Estimate	0.3947	0.5807	1.4061	-0.6840	0.0097	0.0629	0.9126	0.7051
standard error	0.0324	0.0532	0.0419	0.0421	0.0018	0.0156	0.0288	0.1331

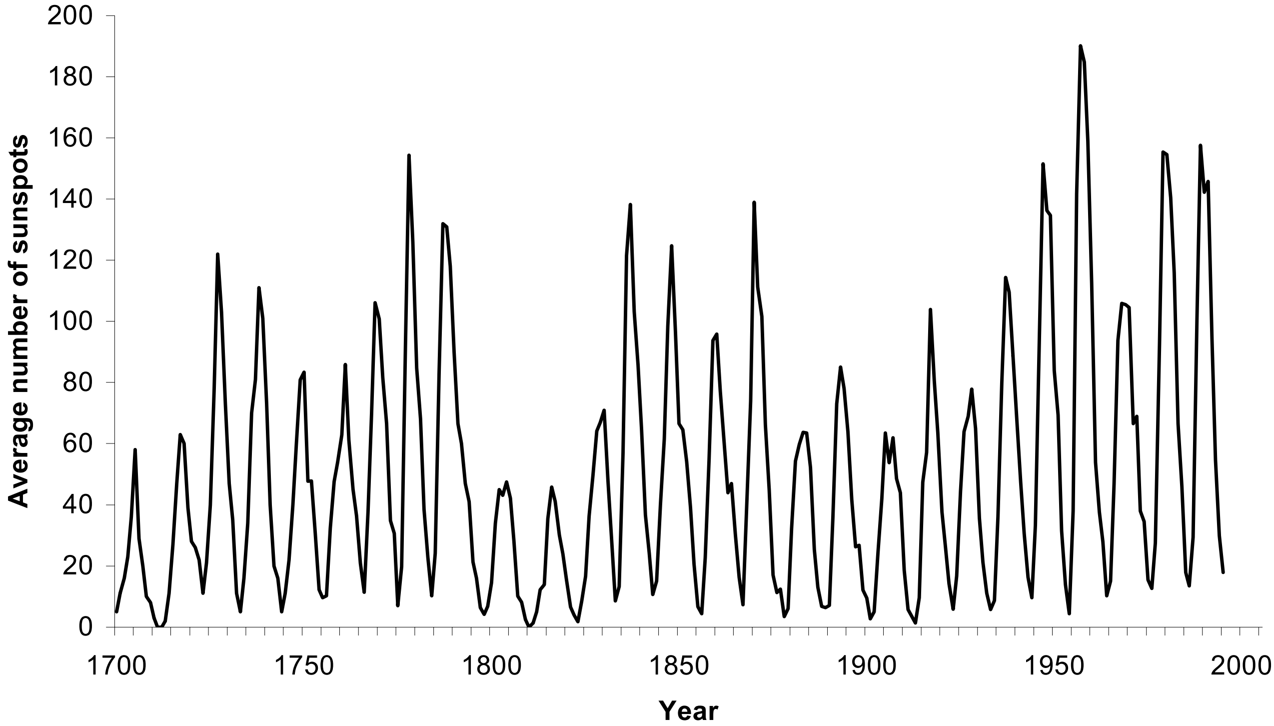

Figure 5: Annual average number of sunspots, 1700 to 1995

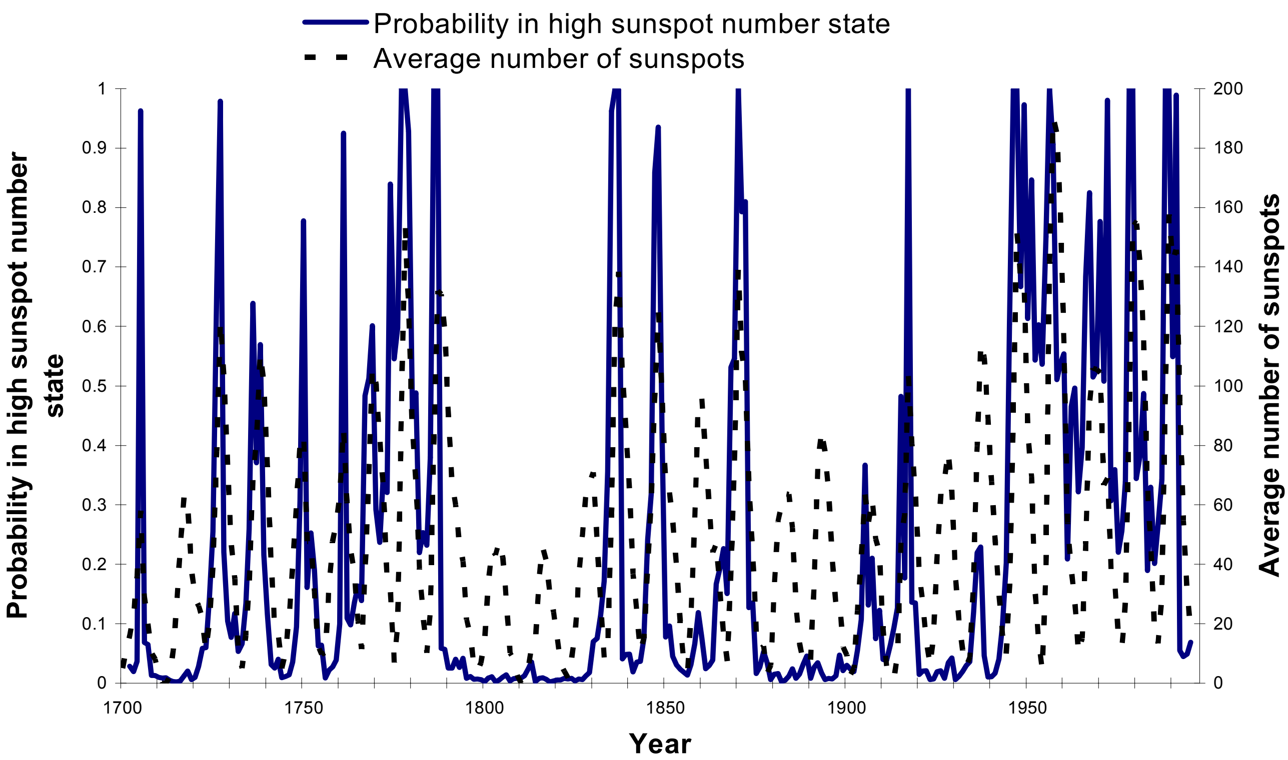

Figure 6: Smoothed probability of being in high sunspot number state (LHS) and number of sunspots (RHS), 1700 to 1995

4 Procedures for aggregating turning point information

So far we have considered two dimensions of the notion of what constitutes decisive change. The first was in terms of whether there was a fall (rise) in a) the level of a specific series; the detrended series; or its k period growth rate. The second dimension of decisive change was in terms of how sustained was the change in direction. The third dimension of what constitutes decisive change relates to how pervasive is that change. If change is pervasive then we would expect to see several series show like turning points at about the same time. That is we would expect to see synchronization of turning points across a range of series that measure economic activity.

Burns and Mitchell (1946, p13) provide a starting point for a definition of what constitutes a set of synchronized cycles with the observation that:

> At an early stage of the investigation we thought it prudent to compare the specific cycles in numerous series. Rough tabulations of specific cycle turns suggested that they clustered around certain months, which usually came in years when business annals reported a recession or revival.

The notion of clusters of turning points is visually appealing but requires careful definition in order to quantify precisely the phenomenon that the eye identifies. Procedures Burns and Mitchell had and their followers at the NBER business cycle dating committee have a long history of interpreting such visual information. Non parametric procedures to extract and codify the rules implicit in the NBER approach were written down by Harding and Pagan (2000) and coded into the computer language GAUSS.

Described in words the algorithm proceeds in the following three stages

1. At date t find the number of months to the nearest peak (trough) for each series. This gives a vector of dimension K. The median of the elements in this vector is then found. The interpretation of this median is that, at time t, it is the median distance to the nearest peak. Designate this item at time t by m_t.

2. Step 1 is done for each t producing $m_t(t = 1, ..., T)$. The series m_t is then examined and, wherever a local minimum is encountered, this is taken to be a candidate for a turning point in the reference cycle.

3. The candidate turning points are then modified in two ways. First, owing to the fact that m_t is discrete, it may be necessary to break ties e.g. m_{J+1} and m_J may be equal and one has to decide whether it is J or $J + 1$ that is the turning point. In this situation the algorithm looks at higher percentiles than the median until a unique local minimum is found. We feel this appeal to clustering in higher order percentiles is a natural way to resolve any non-uniqueness of the local median. Second, turning points may need to be re-combined so that peaks and troughs alternate.

Those procedures were found by Harding and Pagan (2000) to be successful at replicating several judgemental reference cycle chronologies and were found to be useful in extracting a common cycle in industrial production in 12 countries and also a common stock price cycle for Australia, the United Kingdom and the United States. The next section applies this algorithm to obtain the classical, growth and acceleration reference cycles in GDP for the six Euro economies.

5 Obtaining a Euro area reference cycle

5.1 Classical cycle

Earlier we found that the statistical evidence rejected a common classical cycle in the six Euro Area countries but weakly favoured a common cycle in Germany, France, Italy and Austria. The common classical cycle extracted for these four countries is reported in Table 10. Where the dates of peaks and troughs are given along with a measure of the tightness of each cluster of turning points. The later is reported as the median distance in months from the common cycle turning point to each of the specific cycle turning points in the cluster. As can be seen the clusters for the classical cycle are tight with the largest distance being 3 months.

Table 10: Common classical cycle for Germany, France, Italy and Austria

Peaks		Troughs	
Date	Tightness of cluster	Date	Tightness of cluster
1974.9	0	1975.3	0
1980.3	0	1980.12	0
1981.12	3	1982.12	0
1992.3	0	1993.5	1.5

The correlations between the common classical cycle and the specific cycles in each country and for Euro Area GDP are reported in Table 11. The finding that the German cycle is not correlated with the common cycle might seem surprising at first but one should remember that the tests relate to instantaneous synchronization and it is likely that the German cycle either leads or lags the common cycle. Tests for the presence of such leads and lags can be performed but are not undertaken in this paper for reasons of time and space.

Table 11: Selected statistics on the correlation between the common classical cycle for Germany, France, Italy and Austria with Euro Area specific cycles

	GER	FRA	ITA	SPA	NET	AU	EU
Correlation with reference cycle	-0.00	0.80	0.79	0.58	0.37	0.36	0.69
Robust t-statistic	-0.00	2.42	2.00	2.32	1.61	1.42	2.36

5.2 Growth cycle

The tests reported earlier supported the existence of a common Euro Area growth cycle. The common growth cycle extracted for the Euro Area is reported in Table 12. As can be seen from the table the growth cycle turning points are not as tightly clustered as were the classical cycle turning points.

Table 12: Common growth cycle for Germany, France, Italy and Austria

Peaks		Troughs	
Date	Tightness of cluster	Date	Tightness of cluster
1971.4	15	1972.7	13.5
1974.6	3	1975.7	1.5
1977.2	1.5	1977.9	3.0
1980.3	0	1983.1	1.5
1984.2	1.5	1985.5	4.5
1986.7	1.5	1987.3	0.0
1990.3	0	1993.11	1.5
1994.11	4.5		

The correlations between the common cycle and the specific cycles in each country and for Euro Area GDP are reported in Table 13. With the exception of France all of the correlations of the specific cycles with the classical cycle are statistically significant. Although the German cycle is negatively correlated with the common cycle but as discussed above this most likely reflects differences in the lead of the German cycle over the common cycle an issue that requires further investigation.

5.3 Acceleration cycle

As with the growth cycle, the tests reported earlier support the existence of a common Euro Area acceleration cycle. The common cycle extracted from the data is reported in 14 where it is evident that the clusters of turning points are rather tight.

Table 13: Selected statistics on the correlation between the common growth cycle for Germany, France, Italy and Austria with Euro Area specific cycles

	GER	FRA	ITA	SPA	NET	AU	EU
Correlation with reference cycle	-0.35	0.25	0.30	0.41	0.60	0.68	0.45
Robust t-statistic	-2.13	1.74	2.01	2.62	3.43	3.87	2.99

Table 14: Common acceleration cycle for Germany, France, Italy and Austria

Peaks		Troughs	
Date	Tightness of cluster	Date	Tightness of cluster
1972.12	0	1971.8	1.5
1976.6	0	1975.6	0
1979.12	3	1977.9	0
1981.11	1.5	1980.12	0
1983.12	0	1982.9	3
1985.12	3	1984.12	0
1989.7	4.5	1986.12	3
1991.11	1.5	1991.3	3
1994.9	0.0	1992.12	0
		1996.10	1.5

The correlations between the common acceleration cycle and the specific acceleration cycles are generally high and statistically significant. Again the exception is Germany and as mentioned above this most likely reflects differences in the lead of the German economy over the Euro Area an issue that needs further investigation.

6 EuroCoin

EuroCoin is a non parametric coincident index of economic activity that emerged from the work of Forni, Hallin, Lippi, and Reichlin (1999b) and Forni, Hallin, Lippi, and Reichlin (1999a). Figure 7 is taken from the CEPR web page and shows EuroCoin plotted against the raw (seasonally unadjusted) growth rate in Euro Area GDP.

Evidently, EuroCoin is smoother than the raw GDP growth measure but this in part reflects the fact that the GDP data is quarterly and seasonally unadjusted. A more informative comparison would be between EuroCoin and seasonally adjusted GDP interpolated to a monthly basis. EuroCoin finds peaks at 1989.2, 1994.10 and 1997.11 and troughs at 1992.11 and 1995.11 which differ somewhat from the acceleration cycle turning points reported earlier on the basis of my measure of Euro Area GDP. Ultimately one will need to obtain measures of Euro Area GDP on a common basis before being able to compare the common cycle found via EuroCoin with the common cycle found via the non parametric procedures that I discuss above.

Table 15: Selected statistics on the correlation between the common acceleration cycle for Germany, France, Italy and Austria with Euro Area specific cycles

	GER	FRA	ITA	SPA	NET	AU	EU
Correlation with reference cycle	0.03	0.51	0.61	0.57	0.34	0.44	0.32
Robust t-statistic	0.16	3.1	3.56	3.44	2.41	2.91	2.66

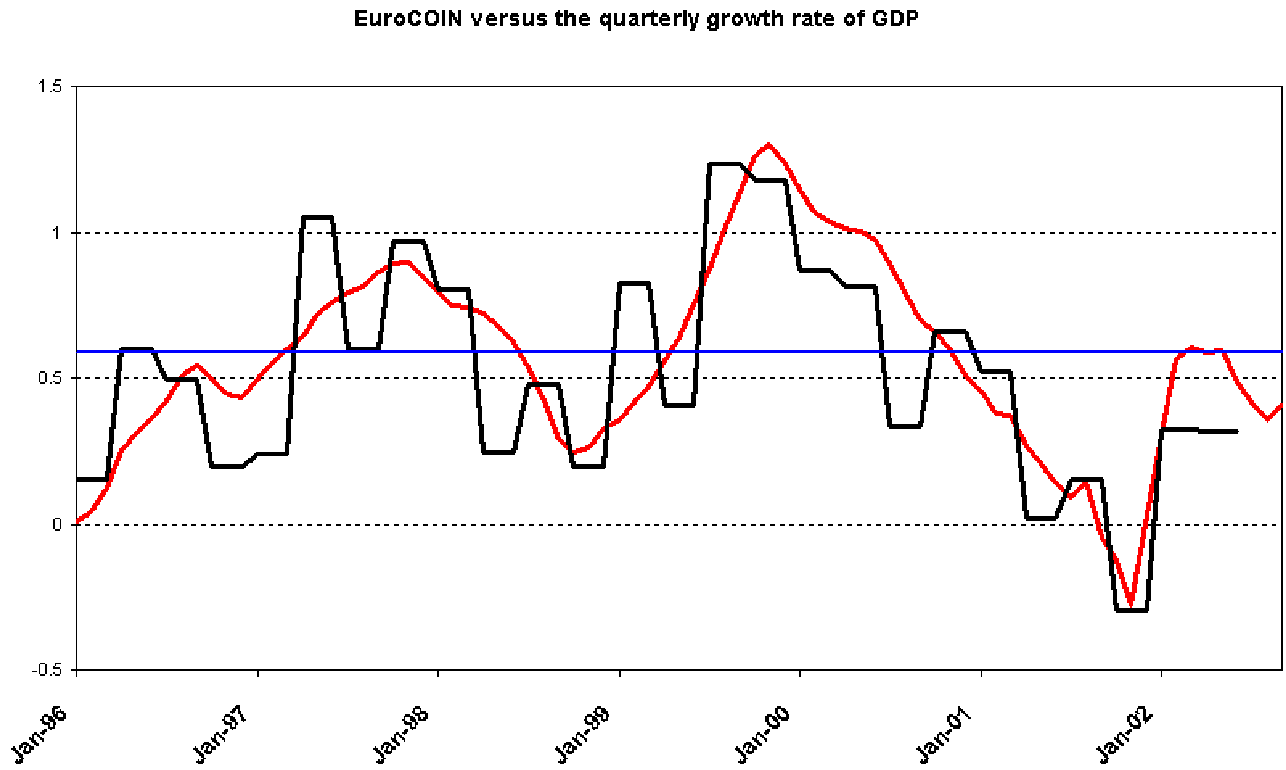

Figure 7: EuroCoin and quarterly GDP

7 Conclusions

Non parametric procedures to locate turning points have been developed and applied to Euro Area data. However, a more comprehensive dataset is required to fully implement the procedures developed in this paper.

References

Boschan, C., and W. W. Ebanks (1978): "The phase-average trend: A new way of measuring economic growth," in *Proceedings of the Business and Economic Statistics Section*, pp. 332–335.

Bry, G., and C. Boschan (1971): *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. NBER.

Christiano, L., and T. Fitzgerald (1998): "The business cycle: It's still a puzzle," *Federal Reserve Bank of Chicago Economic Perspectives*, 4th Quarter, 56–83.

Corbae, D., S. Ouliaris, and P. Phillips (2002): "Band spectral regression with trending data," *Econometrica*, 70(3), 1067–1109.

Diebold, F. X., and G. D. Rudebusch (1999): *Business Cycles: Durations, Dynamics and Forecasting*. Princeton University Press.

Fagan, G., J. Henry, and R. Mestre (2001): "An area wide-model (awm) for the Euro area," Working Paper 42, European Central Bank.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (1999a): "Coincident and leading indicators for the Euro area," mimeo, Universite Libre de Bruxelles, ESCARES.

——— (1999b): "Reference cycles: the NBER methodology revisited," mimeo.

Hamilton, J. D. (1989): "A new approach to the economic analysis of non-stationary times series and the business cycle," *Econometrica*, 57, 357–384.

Harding, D., and A. R. Pagan (2000): "Synchronisation of cycles," The University of Melbourne.

——— (2001): "Extracting, analysing and using cyclical information," Paper presented to the CEPR/Banca d'Italia Conference on Monitoring the Euro Area Business Cycle.

——— (2002a): "A comparison of two business cycle dating methods," *Journal of Economic Dynamics and Control*.

——— (2002b): "Dissecting the cycle: A methodological investigation," *Journal of Monetary Economics*, 49, 365–381.

Hodrick, R. J., and E. C. Prescott (1997): "Post-war US business cycles: A descriptive empirical empirical investigation," *Journal of Money, Credit, and Banking*, 29, 1–16.

Mintz, I. (1969): *Dating postwar business cycles, methods and their application to Western Germany, 1950-1967*. Occasional Paper 107, NBER, New York.

——— (1972): "Dating American growth cycle," in *The Business Cycle Today*, ed. by V. Zarnowitz, pp. 39–88. NBER, New York.

——— (1974): "Dating United States growth cycles," *Explorations in Economic Research*, 1, 1–113.

Pagan, A. R. (2002): "Lecture 1: Extraction of business cycle information," One of three lectures presented at European University Institute, Florence.

Watson, M. W. (1994): "Business-cycle durations and postwar stabilization of the U.S. economy," *American Economic Review*, 84, 24–46.

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Turning Points Detection"

CONSTRUCTING TURNING POINT CHRONOLOGIES WITH MARKOV-SWITCHING VECTOR AUTOREGRESSIVE MODELS: THE EURO-ZONE BUSINESS CYCLE

Hans Martin Krolzig [1]

University of Cambridge

In this paper we advocate a parametric approach to the construction of turning point chronologies for the euro-zone business cycle. In generalization of Hamilton (1989), the Markov-switching vector autoregressive (MS-VAR) model is utilized for the analysis of the business cycle, providing the mechanism for identifying peaks and troughs of the business cycle. Building upon ideas developed in Krolzig (1997a) and Artis, Krolzig, and Toro (2004), the approach for the constructing the turning point chronology consists of (i) modelling the euro-zone business cycle as a single common factor generated by a hidden Markov chain, (ii) fitting a congruent statistical model to the data, (iii) deriving the conditional probabilities of the regimes 'expansion' and 'recession' from the estimated model, (iv) classifying each point in time to the regime with the highest probability, and (v) dating the turning points of the business cycle. The MS-VAR also provides measures of uncertainty associated with the turning point chronology, facilitates real-time detection of business cycle transitions, and offers a well-developed theory for the prediction of the business cycle. Examining the properties of the proposed dating procedure, we show that aggregation and model misspecification can severely hamper the detection of business cycle turning points. In the empirical part, the MS-VAR approach is applied to three multi-country data sets consisting of real GDP and industrial production growth rates of 12 euro-zone countries (in total) over the period from 1973 to 2002. Although the empirical models are found to be sensitive to data quality (with seasonal adjustment, outlier correction and smoothing being all influential), the estimated models were found useful for the assessment of business cycle synchronization among the EMU member states, and for the construction of a turning point chronology of the euro-zone business cycle.

Keywords: Markov switching; Business cycles; Turning points; Monetary union; European Union.

JEL Classification: E32; F43; F47; C32.

[1] I am grateful to Gian Luigi Mazzi for the invitation to contribute this paper to the colloquium and for providing the Eurostat data set. Address of the author: Hans-Martin Krolzig, Department of Economics, University of Oxford, Manor Road Building, Oxford OX1 3UQ, UK.

1 Introduction

The creation of the euro zone with begin of Stage Three of the Economic and Monetary Union (EMU) on 1 January 1999 has raised several important issues. Among them, one of paramount relevance concerns the economic convergence of the EMU member states. A strong indicator for convergence is the existence of a common euro-zone business cycle. Furthermore, business cycle synchronization constitutes a pre-condition for symmetry of the effects of monetary policy in the euro zone. The presence of a common euro-zone business cycle implies the coincidence of business cycle turning points in the countries of the euro-zone. The detection of its turning points of the common cycle is, therefore, crucial for monetary policy making in the union.

There is a huge literature evaluating the degree of business cycle synchronization in the EMU. Among others, Artis, Krolzig, and Toro (2004) conclude that '*the conception of a common or "European" business cycle is an intelligible one*'. Building upon that result, this paper deals with the construction of a turning point chronology for the euro-zone business cycle. A parametric approach is advocated for doing so, which consists of first fitting a congruent statistical model to the data, and then utilizing the estimated coefficients for the classification of observations as *'recession'* or *'expansion'*; the implied dating of the turning points finally constitutes the business cycle chronology. The approach is based on the Markov-switching time series model innovated by Hamilton (1989) in his analysis of the US business cycle and generalized in Krolzig (1997b) to a Markov-switching vector autoregressive model, which will be used for the modelling and dating of the euro-zone business cycle. In the framework considered here, the euro-zone business cycle is defined as the common factor of the macroeconomic fluctuations in the countries of the euro-zone, and it is assumed that the transition of its phases are generated by an unobservable stochastic process, which is a hidden Markov chain.

The paper proceeds as follows. First, in § 2 we discuss the Markov-switching time series model and its application to the modelling of the business cycle. Commencing from the univariate Hamilton (1989) model of the US Business Cycle, the Markov-switching vector autoregressive (VAR) model of the business cycle is presented. Then, § 3 considers the construction of turning point chronologies with Markov-switching models. We present the regime inference in Markov-switching VAR models and discuss how the derived regime probabilities can in turn be used to date the turning points of the business cycle, finally producing the turning point chronology. We also point out three distinctive advantages of the MS-VAR approach: the provision of an uncertainty measure as a by-product of the regime classification, the possibility of real-time detecting of turning points, and a well-developed theory of the prediction of the business cycle. In § 4 the properties of the proposed regime classification rule are evaluated for univariate and multivariate Markov-switching models. We analyze the effects of business cycle synchronization and regime-dependent heteroscedasticity on the classification procedure, and show that aggregation obstructs the detection of turning points, when the variable are driven by a common business cycle. Section 5 presents an empirical analysis of the euro-zone business cycle. Starting with a review of the literature featuring Markov-switching studies of the European business cycle, we discuss the empirical modelling approach used in this paper, and present results for three Markov-switching VAR models of the euro-zone business cycle using quarterly OECD GDP data (1973-2002), OECD industrial production data (1974-2002), and Eurostat GDP data (1980-2002), respectively. After having established evidence for business cycle synchronization in the euro area, the estimated models are used for the construction of turning point chronologies of the euro-zone business cycle. Finally § 6 concludes.

2 Business Cycle Modelling with Markov-switching Vector Autoregressions

Recent theoretical and empirical business cycle research has revived interest in the co-movement of macroeconomic time series and the regime-switching nature of macroeconomic activity. The general idea behind regime-switching models of the business cycle is that the parameters of a time series model of some macroeconomic variables depend upon a stochastic, unobservable regime variable $s_t \in \{1, \ldots, M\}$ which represents the state of business cycle. The number of regimes, M, is often assumed to be two reflecting economic expansions and contractions.

The Markov-switching autoregressive time series model has emerged as a leading approach for the detection and dating of business cycle turning points. Since Hamilton's (1989) application of this technique to measure the US business cycle, there has been a number of subsequent extensions and refinements (see Krolzig (1997b) for an overview). We first present the original contribution of Hamilton, then consider the multivariate generalization, and finally discuss its application to the business cycle.

2.1 Hamilton's Model of the US Business Cycle

In the original contribution of Hamilton (1989), contractions and expansions are modelled as switching regimes of the stochastic process generating the growth rate of real output Δy_t:

$$\Delta y_t - \mu(s_t) = \alpha_1 \left(\Delta y_{t-1} - \mu(s_{t-1})\right) + \ldots + \alpha_4 \left(\Delta y_{t-4} - \mu(s_{t-4})\right) + u_t. \tag{1}$$

In (1), the two regimes are associated with different conditional distributions of the growth rate of real output, where the mean growth rate μ depends on the state or 'regime', s_t. For a meaningful business cycle model, the mean growth rate should be negative in the first regime, $\mu_1 < 0$ ('*recession*'), and positive in the second, $\mu_2 > 0$ ('*expansion*'). The variance of the disturbance term, $u_t \sim \mathsf{NID}(0, \sigma^2)$, is assumed to be the same in both regimes.

The stochastic process generating the unobservable regimes is an ergodic Markov chain defined by the constant transition probabilities:

$$\begin{aligned} p_{12} &= \Pr(\text{ recession in } t \mid \text{expansion in } t-1), \\ p_{21} &= \Pr(\text{ expansion in } t \mid \text{recession in } t-1), \end{aligned}$$

The transition probabilities have to be estimated together with the parameters of equation (1).

By inferring the probabilities of the unobserved regimes conditional on the available information set, it is possible to reconstruct the regimes and to date the turning points of the business cycle. Hamilton was able to show that his model is able to track the NBER reference cycle very closely. This suggests that the Markov-switching approach should also be useful for the modelling and dating of the euro-zone business cycle.

In the literature (see the review in § 5.1), univariate Markov-switching models like the one in (1) have been applied to aggregated euro-zone data (with mixing results). However, it is important to note that, by definition, univariate Markov-switching models as proposed by Hamilton (1989) are only able to capture some of the stylized facts of the business cycle. They can represent the non-linearity or asymmetry stressed in some part of the literature, but obviously, they are unable to reflect the idea of co-movement among time economic series. This is particularly important in the context of the euro-zone since there is a genuine interest in assessing the synchronization of the business cycles among the euro-zone countries.

In this paper, we think of the euro-zone business cycle as a common cyclical feature of the macroeconomic dynamics in the EMU member states. It is a common factor, which shares all the properties of the regime variable in (1) by being generated by a hidden Markov chain. We shall therefore utilize multivariate generalizations of Hamilton's approach for the modelling of the euro-zone business cycle using real GDP and industrial production time series for the countries of the euro-zone (to be presented in § 5). Identifying the transitions of the regime variable will give us the turning points of the common cycle component of the EMU member states leading towards a chronology of the euro-zone business cycle. Next, we introduce the statistical model: the *Markov-switching Vector Autoregression*. Its interpretation as a business cycle model follows in § 2.3).

2.2 Markov-switching Vector Autoregressive Models

Like other regime-switching models, the Markov-switching vector autoregressive (MS-VAR) model is a vector autoregressive process of the *observed* time series vector $\mathbf{y}_t = (y_{1t}, \dots, y_{Kt})'$, whose parameters are, at least partly, unconditionally time-varying but constant when conditioned on an *unobservable* discrete regime variable $s_t \in \{1, \dots, M\}$:

$$\mathbf{y}_t - \boldsymbol{\mu}(s_t) = \mathbf{A}_1(s_t)\,(\mathbf{y}_{t-1} - \boldsymbol{\mu}(s_{t-1})) + \ldots + \mathbf{A}_p(s_t)\,(\mathbf{y}_{t-p} - \boldsymbol{\mu}(s_{t-p})) + \mathbf{u}_t, \tag{2}$$

where $\mathbf{u}_t$ is a Gaussian error term conditioned on s_t: $\mathbf{u}_t|s_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma}(s_t))$. A p-th order, M-state Markov-switching vector autoregressive model of a Kdimensional time series vector is denoted MS(M)–VAR(p). The parameter matrix shift functions $\boldsymbol{\mu}(s_t)$, $\mathbf{A}_1(s_t), \dots, \mathbf{A}_p(s_t)$ and $\boldsymbol{\Sigma}(s_t)$ describe the dependence of the VAR parameters $\boldsymbol{\mu}$, $\mathbf{A}_1, \dots, \mathbf{A}_p$ and $\boldsymbol{\Sigma}$ on the regime variable s_t as in:

$$\boldsymbol{\mu}(s_t) = \begin{cases} \boldsymbol{\mu}_1 = (\mu_{11}, \dots, \mu_{K1})' & \text{if } s_t = 1, \\ \vdots & \\ \boldsymbol{\mu}_M = (\mu_{1M}, \dots, \mu_{KM})' & \text{if } s_t = M. \end{cases} \tag{3}$$

The decisive characteristic of a Markov-switching model is that the unobservable realizations of the regime $s_t \in \{1, \dots, M\}$ are generated by a discrete time, discrete state Markov stochastic process, which is defined by its transition probabilities:

$$p_{ij} = \Pr(s_{t+1} = j | s_t = i), \quad \sum_{j=1}^{M} p_{ij} = 1 \quad \text{for all } i, j \in \{1, \dots, M\}. \tag{4}$$

It is usually assumed that the Markov process is irreducible and ergodic.

In the model (2) there is an immediate one–time jump in the process mean after a change in the regime. Frequently, it is more plausible to assume that the mean smoothly approaches a new level after the transition from one state to another. In such a situation, the following model with a regime-dependent intercept term $\boldsymbol{\nu}(s_t)$ could be used:

$$\mathbf{y}_t = \boldsymbol{\nu}(s_t) + \mathbf{A}_1(s_t)\mathbf{y}_{t-1} + \ldots + \mathbf{A}_p(s_t)\mathbf{y}_{t-p} + \mathbf{u}_t. \tag{5}$$

In contrast to linear time-invariant VAR models, the mean adjusted form (2) and the intercept form (5) are not equivalent. They imply different dynamic adjustments of the observed variables after a change in regime. While a permanent regime shift in the mean $\boldsymbol{\mu}(s_t)$ causes an immediate jump of the observed time series vector onto its new level, the dynamic response to a once-and-for-all regime shift in the intercept term $\boldsymbol{\nu}(s_t)$ is identical to an equivalent shock in the white noise series $\mathbf{u}_t$.

The MS-VAR model allows for a great variety of specifications. In principle, it would be possible to alternate the specification of the regime generating process (i) by making all parameters time-varying and (ii) by introducing separate regimes for each equation (or even for each parameter). But, this is not a practicable solution as the number of parameters of the Markov chain grows quadratic in the number of regimes and coincidently shrinks the number of observations usable for the estimation of the regime-dependent parameters. For these reasons, we prefer to impose the assumption of a common regime variable generating simultaneous regime shifts in all equations of the system, *i.e.* identical break-points for all variables and parameters, and also to restrict the number of regime-dependent parameters.

In empirical research, only some parameters will be conditioned on the state of the Markov chain while the other parameters will be regime invariant. In order to establish a unique notation for each model, we specify with the general MS(M) term the regime-dependent parameters:

M Markov-switching *mean* ,
I Markov-switching *intercept*,
A Markov-switching *autoregressive parameters*,
H Markov-switching *heteroscedasticity*.

The MS-VAR model provides a very flexible framework which allows for heteroscedasticity, occasional shifts, reversing trends, and forecasts performed in a non-linear manner. The model can be easily extended for the presence of cointegration (see Krolzig (1996), for the statistical analysis of these systems and Krolzig (2001a), for applications to the business cycle). In business cycle analysis, the focus is on models where the mean (MSM(M)–VAR(p) models) or the intercept term (MSI(M)–VAR(p) models) are subject to occasional discrete shifts creating episodes of economic expansion versus contraction or stagnation. Regime-dependent covariance structures are often required to model changes in the correlation structure or the considered as additional features.

2.3 MS-VAR Models of the Business Cycle

We start by presenting a simple, multivariate extension of Hamilton's original model in (1). Since we are interested in modelling the business cycle, only shifts in the vector of mean growth rates $\boldsymbol{\mu}(s_t)$ are considered:

$$\Delta \mathbf{y}_t - \boldsymbol{\mu}(s_t) = A_1 \left(\Delta \mathbf{y}_{t-1} - \boldsymbol{\mu}(s_{t-1})\right) + \ldots + A_p \left(\Delta \mathbf{y}_{t-p} - \boldsymbol{\mu}(s_{t-p})\right) + \mathbf{u}_t, \tag{6}$$

where the error term $\mathbf{u}_t$ is again Gaussian: $\mathbf{u}_t | s_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma})$. The MSM($M$)-VAR($p$) characterizes business cycles as common regime shifts in the stochastic process of some macroeconomic time series. In the context of this paper, those are the growth rates of macroeconomic activity in the euro-zone countries.

MS-VAR models are by no means restricted to the presence of two regimes as assumed by Hamilton (1989). In practice, two-regime models exhibiting episodes of '*recession*' and '*expansion*' may not congruent since, for example, structural change may require to allow for multiple expansionary regimes such as '*high growth*' and '*low growth*' (see, *inter alia*, Krolzig (2001a)). Also, note that without further restrictions on the matrix of transition probabilities, for $M > 2$, the model does not impose a sequential ordering of the regimes (*e.g.*, a rigid phase structure of business cycle). In (6) each regime, say $s_t = m$, is associated with a particular mean vector $\boldsymbol{\mu}_m = (\mu_{1m}, \ldots, \mu_{Km})'$ such that $\mathsf{E}\left[\Delta \mathbf{y} | s_t = m\right] = \boldsymbol{\mu}_m$. As a normalization, the regimes are ordered with increasing mean, *i.e.* $\|\boldsymbol{\mu}_1\| < \|\boldsymbol{\mu}_2\|$. Identification requires that for all $m \in \{1, \ldots, M\}$ there exist at least one $k \in \{1, \ldots, K\}$ such that $\mu_{km} \neq \mu_{1i}$ for all $i \neq m$. Assume, for instance, that the mean growth rates are identical for all countries, $\mu_{km} = \mu_m$ for all k and m. Then, in the case of $M = 2$: $\mu_1 < 0 < \mu_2$ implies that latter regime represents the expansionary phase and the former the recessionary phase of the *classical* business cycle. Suppose now that $0 < \mu_1 < \mu_2$, then the two regimes correspond to the phases of the *growth* cycle. When the growth rates vary across countries, the distinction between classical and growth cycles can be more difficult. In general, we will require that at least some country-specific mean growth rates are negative during recessions. A stronger condition may demand that a well defined index or aggregate of $\mathbf{y}_t$ contracts during recessions: $E[y_t^a | s_t = 1] = \mathbf{a}'\boldsymbol{\mu}_1 < 0$, where $y_t^a = \mathbf{a}'\mathbf{y}_t$ with $\mathbf{a} > 0$ and $\mathbf{1}'\mathbf{a} = 1$. Similar criteria can be applied to distinguish the phases of the growth cycle, *i.e.* $\mathbf{a}'\boldsymbol{\mu}_2 > \mathbf{a}'\boldsymbol{\mu}_1 > 0$.

Despite its simplicity, the model is able to capture the non-linear, regime-switching, and the common factor structures of the business cycle simultaneously. The design of the model allows the identification and analysis of the latent euro-zone cycle, while restricting the propagation of country-specific shocks to the vector autoregression. In other words, the hidden Markov process is not informative about the propagation of business cycles across countries. Modelling the '*business cycle*' as a common factor implies that the countries share the same turning points. However, this does not presuppose that all countries are synchronized with the common cycle, *i.e.* it is feasible that some countries – but not all – are unaffected by shifts in regime: $\mu_{km} = \mu_k$ for all m. Clearly, the model would not be identified if this were true for all k. A more flexible approach can be found in Phillips (1991) and Hamilton and Lin (1996), where some variables might precede others in their cycle or different variables follow completely different Markov processes. Given the relatively large number of countries considered and our special interest in the phenomenon of a *common* euro-zone cycle, introducing separate Markov processes for each country is not very appealing within the modelling strategy outlined here. In other instances, however, it may be appropriate to allow more than one state variable, such that each variable may respond to a specific state variable.[2] It is worth pointing out that if (6) were formulated

[2]Considering two-country two-regime models of monthly growth rates of industrial production in the US,

as an MSI(M)–VAR(p) model, regime shifts would cause lagged responses of the variables in the system, so that some countries would react faster to business cycle transitions than others.

In many regards, Markov-switching VAR models can be considered as rather straightforward multivariate generalizations of the univariate Markov-switching model originally proposed by Hamilton (1989) for the US business cycle. This also helps to explain some of their distinctive attributes: they are foremost statistical models designed to learn the features of the business cycle from the data. As statistical representations of the business cycle, they require minimal theoretical prior information about the relevant macroeconomic structures. The common factor hypothesis is such an identifying assumption: one regime variable drives the business cycle transitions in the variables of the system. Given the minimal input of identifying assumptions regarding the regime generating process, the interpretation of its realization as the '*business cycle*' must be based on similarities with a reference cycle. This may be called an '*ex-post identification of the business cycle*'. The other important assumption made in Markov-switching models says that the latent regime variable is generated by a Markov chain. This allows the '*reconstruction*' of the evolution of the regime (*e.g.*, as the time path of the '*probability of being in a recession*') by solving a *signal-extraction problem.* If the business cycle is a common feature of the vector of economic time series, then modelling the system does not only reflect the basic definition of the business cycle, but also enhances the extraction of the common 'business cycle' component from the variables. In the next two sections, we explore this issue further by discussing the regime inference in MS-VAR models, its exploitation for the construction of a turning point chronology as well as the properties of resulting regime classification.

3 Constructing the Turning Point Chronology

For an estimated MS-VAR model, the construction of a business cycle chronology consists of three major steps:

i. *Regime inference*: identification of the unobserved regime variable s_t from the information set $\mathbf{Y}_T$

- Filtering: $\Pr(s_t|\mathbf{Y}_t)$;
- Smoothing: $\Pr(s_t|\mathbf{Y}_T)$;
- Predicting: $\Pr(s_{t+h}|\mathbf{Y}_T)$.

ii. *Construction of the business cycle chronology*

- Regime classification using a classification rule based on $\Pr(s_t|\mathbf{Y}_T)$: $\hat{s}_t$;
- Dating the turning points.

iii. *Control of the regime classification*

the UK, Germany and Japan, Phillips (1991) assumes separate Markov chains for the business cycle of each country. Interestingly, in none of the estimated models, the null hypothesis of perfectly correlated regime shifts could be rejected which supports the presumptions made here.

- Uncertainty: $\sum_{s_t \neq \hat{s}_t} \Pr(s_t|\mathbf{Y}_T)$;
- Reliability:$\|\Pr(s_t|\mathbf{Y}_T) - \Pr(s_t|\mathbf{Y}_t)\|$;
- Robustness;
- Parameter constancy.

iv. *Implications for real-time turning point detection.*

3.1 Regime Inference in MS-VAR Models

The framework for the statistical analysis of MS-VAR models is the state-space form, where the VAR of $\mathbf{y}_t$ conditional on the state variable s_t in equation (6) forms the measurement equation, and the hidden Markov chain in equation (4) is the transition equation of the state variable (see Krolzig (1997b), ch.2, for more details). In the following, we only give a brief introduction to the calculation of the filtered and smoothed regime probabilities, which are essential for the dating of the business cycle with MS-VAR models.

3.1.1 Filtering

The filtering algorithm is usually associated with Hamilton (1989), but it can be traced back to earlier statistics literature including Baum and Petrie (1966) and Lindgren (1978). The aim of the filter is to infer the probability distribution of the unobserved regime variable s_t given the currently available information set $\mathbf{Y}_t$.

Let $\mathsf{p}(\mathbf{y}_t|s_t, \mathbf{Y}_{t-1})$ denote the probability density of observing $\mathbf{y}_t$ conditional on the regime variable s_t, and $\Pr(s_t|s_{t-1})$ the transition probability from regime s_{t-1} to s_t. By invoking the law of Bayes, the (posterior) probability $\Pr(s_t|\mathbf{Y}_t)$ of the regime variable s_t conditional on all available information at time t is given by:

$$\Pr(s_t|\mathbf{Y}_t) = \Pr(s_t|\mathbf{y}_t, \mathbf{Y}_{t-1}) = \frac{\mathsf{p}(\mathbf{y}_t|s_t, \mathbf{Y}_{t-1}) \Pr(s_t|\mathbf{Y}_{t-1})}{\mathsf{p}(\mathbf{y}_t|\mathbf{Y}_{t-1})}, \tag{7}$$

where the predicted regime probability $\Pr(s_t|\mathbf{Y}_{t-1})$ is the *prior* probability of regime s_t given the information set of the previous period and $\mathsf{p}(\mathbf{y}_t|\mathbf{Y}_{t-1})$ is the *marginal* density of $\mathbf{y}_t$ conditional on the information set $\mathbf{Y}_{t-1}$. The predicted regime probabilities can be calculated as:

$$\Pr(s_t = j|\mathbf{Y}_{t-1}) = \sum_{i=1}^{M} \sum_{j=1}^{M} p_{ij} \Pr(s_{t-1} = i|\mathbf{Y}_{t-1}),$$

which in turn are used in the derivation of the predicted density of $\mathbf{y}_t$:

$$\mathsf{p}(\mathbf{y}_t|\mathbf{Y}_{t-1}) = \sum_{j=1}^{M} \mathsf{p}\left(\mathbf{y}_t, s_t = j \mid \mathbf{Y}_{t-1}\right) = \sum_{s_t} \mathsf{p}(\mathbf{y}_t|s_t, \mathbf{Y}_{t-1}) \Pr(s_t|\mathbf{Y}_{t-1}). \tag{8}$$

Using (7) the filtered regime probabilities for a sample $\mathbf{Y}_T = (\mathbf{y}_T, \ldots, \mathbf{y}_1)$ can be calculated by a forward recursion for $t = 1, \ldots, T$ initialized by some estimate of the initial value s_0 of the regime variable. Suppose, for example, $p = 0$ and $M = 2$ where one regime represents '*recession*' and the other '*expansion*', then the filter recursion in (7) can be simplified to the following odds ratio of the regimes:

$$\frac{\Pr(\text{'}recession\text{' at time } t|\mathbf{Y}_t)}{\Pr(\text{'}expansion\text{' at time } t|\mathbf{Y}_t)} = \frac{\mathsf{p}(\mathbf{y}_t|\text{'}recession\text{'})}{\mathsf{p}(\mathbf{y}_t|\text{'}expansion\text{'})} \frac{\Pr(\text{'}recession\text{' at time } t|\mathbf{Y}_{t-1})}{\Pr(\text{'}expansion\text{' at time } t|\mathbf{Y}_{t-1})}.$$

The probability of a recession given all available information at time t depends (i) on the relative likelihood of observing $\mathbf{y}_t$ in a recession versus in an expansion and (ii) on the relative likelihood of a recession given the history of $\mathbf{y}_t$, namely the information set $\mathbf{Y}_{t-1}$ available in the previous period.

3.1.2 Smoothing

The filter recursions deliver estimates for s_t, with $t = 1, \ldots, T$, based on the available information up to time point t. This is a limited information technique, as we have observations up to $t = T$. Thus, the regime inference can be improved further by using future observations of $\mathbf{y}_t$, in which case the resulting regime probabilities, $\Pr(s_t|\mathbf{Y}_s)$ with $s > t$, are called '*smoothed*'. The smoothing algorithm gives the best estimate of the unobservable state at any point within the sample.

In the following, we discuss how the sample information $\mathbf{Y}_{t+1.T} = (\mathbf{y}'_{t+1}, \ldots, \mathbf{y}'_T)'$ neglected so far is incorporated into the full-sample inference about the unobserved regime s_t. The smoothing algorithm proposed by Kim (1994) may be interpreted as a backward filter that starts at the end point, $t = T$, of the previously applied filter. The full–sample smoothed inferences $\Pr(s_t|\mathbf{Y}_T)$ are found by starting from the last output of the filter $\Pr(s_T|\mathbf{Y}_T)$, and then by iterating backward from $t = T - 1$ to $t = 1$.

The algorithm exploits the identity:

$$\begin{aligned} \Pr(s_t|\mathbf{Y}_T) &= \sum_{s_{t+1}} \Pr(s_t, s_{t+1}|\mathbf{Y}_T) \\ &= \sum_{s_{t+1}} \Pr(s_t|s_{t+1}, \mathbf{Y}_T)\Pr(s_{t+1}|\mathbf{Y}_T). \end{aligned} \tag{9}$$

For pure VAR models with Markovian parameter shifts, the probability laws for y_t and s_{t+1} depend only on the current state s_t, but not on its history. Thus, we have:

$$\begin{aligned} \Pr(s_t|s_{t+1}, \mathbf{Y}_T) &\equiv \Pr(s_t|s_{t+1}, \mathbf{Y}_t, \mathbf{Y}_{t+1.T}) \\ &= \frac{\mathsf{p}(\mathbf{Y}_{t+1.T}|s_t, s_{t+1}, \mathbf{Y}_t)\Pr(s_t|s_{t+1}, \mathbf{Y}_t)}{\mathsf{p}(\mathbf{Y}_{t+1.T}|s_{t+1}, \mathbf{Y}_t)} \\ &= \Pr(s_t|s_{t+1}, \mathbf{Y}_t). \end{aligned}$$

It is, therefore, possible to calculate the smoothed probabilities $\Pr(s_t|\mathbf{Y}_T)$ by getting the last term from the previous iteration of the smoothing algorithm, $\Pr(s_{t+1}|\mathbf{Y}_T)$, while it can be shown that the first term can be derived from the filtered probabilities $\Pr(s_t|\mathbf{Y}_t)$:

$$\begin{aligned} \Pr(s_t|s_{t+1}, \mathbf{Y}_t) &= \frac{\Pr(s_{t+1}|s_t, \mathbf{Y}_t)\Pr(s_t|\mathbf{Y}_t)}{\Pr(s_{t+1}|\mathbf{Y}_t)} \\ &= \frac{\Pr(s_{t+1}|s_t)\Pr(s_t|\mathbf{Y}_t)}{\Pr(s_{t+1}|\mathbf{Y}_t)}. \end{aligned} \tag{10}$$

The recursion is initialized with the final filtered probability vector $\Pr(s_t|\mathbf{Y}_T)$:

$$\Pr(s_t|\mathbf{Y}_T) = \sum_{s_{t+1}=1}^{M} \frac{\Pr(s_{t+1}|s_t)\Pr(s_t|\mathbf{Y}_t)}{\Pr(s_{t+1}|\mathbf{Y}_t)} \Pr(s_{t+1}|\mathbf{Y}_T). \tag{11}$$

Recursion (11) describes how the additional information $\mathbf{Y}_{t+1.T}$ is used in an efficient way to improve the inference on the unobserved state s_t:

$$\frac{\Pr(s_t|\mathbf{Y}_T)}{\Pr(s_t|\mathbf{Y}_t)} = \sum_{s_{t+1}=1}^{M} \Pr(s_{t+1}|s_t) \frac{\Pr(s_{t+1}|\mathbf{Y}_T)}{\Pr(s_{t+1}|\mathbf{Y}_t)}. \tag{12}$$

If at time $t+1$ the full information inference, $\Pr(s_{t+1}|\mathbf{Y}_T)$, coincides with its prediction, $\Pr(s_{t+1}|\mathbf{Y}_t)$, then knowing the future of $\mathbf{y}_t$ does not help to improve the regime inference at time t. Hence the filtering solution $\Pr(s_t|\mathbf{Y}_t)$ is not updated: $\Pr(s_t|\mathbf{Y}_T) = \Pr(s_t|\mathbf{Y}_t)$.

In the case of two regimes, say '*recession*' ($s_t = 1$) and '*expansion*' ($s_t = 2$) , we have that:

$$\begin{aligned}
&\frac{\Pr(\text{`\textit{recession}' at time } t|\mathbf{Y}_T)}{\Pr(\text{`\textit{recession}' at time } t|\mathbf{Y}_t)} \\
&= p_{11}\frac{\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_T)}{\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_t)} + p_{12}\frac{\Pr(\text{`\textit{expansion}' at time } t+1|\mathbf{Y}_T)}{\Pr(\text{`\textit{expansion}' at time } t+1|\mathbf{Y}_t)} \\
&= p_{11}\frac{\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_T)}{\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_t)} + (1-p_{11})\frac{1-\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_T)}{1-\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_t)}.
\end{aligned} \tag{13}$$

It can be shown that iff $p_{11} > p_{21}$, *i.e.*, whenever the probability of staying in a recession is greater than the probability of moving from an expansion to a recession, $\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_t)$ is a monotonically increasing function in $\Pr(\text{`\textit{recession}' at time } t+1|\mathbf{Y}_t)$. In other words, if new information will make the researcher believe that the probability of a recession in $t+1$ is greater than expected at time t, she will update her beliefs about the past by increasing the probability that the economy has already been in a recession at time t. Finally, note that the difference between the two measures indicates the amount of revisions in the regime classification, when new observations become available.

3.2 Constructing the Turning Point Chronology

3.2.1 Classification of the Regimes

In Markov-switching models, the classification of the regimes and the dating of the business cycle amounts to assigning every observation $\mathbf{y}_t$ to a regime $\hat{s}_t \in \{1, \ldots, M\}$. The rule that is applied here maps $\mathbf{y}_t$ to the regime with the highest smoothed probability:

$$\hat{s}_t = \arg\max_{1,\ldots,M} \Pr(s_t \mid \mathbf{Y}_T) \tag{14}$$

At each point in time, the smoothed regime probabilities are calculated as described in §3.1, *i.e.* the inference is made using the whole set of data points, and each observation is matched with the most likely regime.

In the simplest case of two regimes, the classification rule proposed rule in (14) simplifies to assigning the observation to the first regime if $\Pr(s_t = 1 \mid \mathbf{Y}_T) > 0.5$ and to the second if $\Pr(s_t = 1 \mid \mathbf{Y}_T) < 0.5$, which was proposed in Hamilton (1989). For the purpose of business cycle analysis, this rule might also be applied to multi-regime models ($M > 2$) when, for instance, regime 1 captures '*recessions*' and the remaining regimes ($s_t = 2, \ldots, M$) can be aggregated to the superstate '*expansion*' with each regime representing different phases of an expansion.

3.2.2 Dating the Business Cycle Turning Points

The regime classification procedure allows dating the turning points of the business cycle by identifying the transitions of the regime variable. Suppose again that there are two-regimes: '*recession*' ($s_t = 1$) and '*expansion*' ($s_t = 2$). Then, the peak date denotes the period t just before the beginning of a recession, i.e. Pr('*recession*' at time $t \mid \mathbf{Y}_T) < 0.5$ and Pr('*recession*' at time $t+1 \mid \mathbf{Y}_T) > 0.5$; the trough is the last period of the recession:

$$\begin{aligned} \text{Peak at } t = \tau \quad &\Longleftrightarrow \quad \hat{s}_\tau = 2 \text{ and } \hat{s}_{\tau+1} = 1 \\ \text{Trough at } t = \tau \quad &\Longleftrightarrow \quad \hat{s}_\tau = 1 \text{ and } \hat{s}_{\tau+1} = 2. \end{aligned}$$

It would also be feasible to impose the additional condition that each recession and expansion should prevail for at least two quarters, which would add a certain smoothness to the regime classification. A chronology of the turning points of the euro-zone business cycle will be discussed in § 5.5 (see Tables 9 and 10).

3.3 Quantifying the Classification Uncertainty

An important advantage of the MS-VAR model is its ability not only to classify observations, but also to quantify the uncertainty associated with the produced business cycle chronology. If we attach the observation at time t to the regime $\hat{s}_t$ according to rule (14), then an appropriate measurement of the classification uncertainty is given by:

$$\frac{M}{M-1} \sum_{i \neq \hat{s}_t} \Pr(s_t = i | \mathbf{Y}_T),$$

where $1 - 1/M$ is the maximal uncertainty when all regimes $m = 1, \ldots, M$ are possible with the same probability $1/M$. Hence, the proposed measurement is bounded between 0 and 1. Obviously, we get for $M = 2$, that the probability of a wrong classification, which is given by the complementary probability, is normalized to $2 \Pr(s_t \neq \hat{s}_t | \mathbf{Y}_T)$.

A different measure of uncertainty is given by the difference of the filtered and smoothed probabilities of a recession, which reports the update of the regime inference as new information becomes available. This is crucial for the reliability of real-time detection of business cycle transitions with MS-VAR models, which we discuss next.

3.4 Real-Time Detecting of Turning Points

The proposed approach supports the identification of business cycle turning points without a lag. The filtered regime probability $\Pr(s_T|\mathbf{Y}_T)$, which coincides with the smoothed probability of s_T for the sample $\mathbf{Y}_T$, constitutes the optimal regime inference for given data set. When additional data become available, the regime inference will be updated by the smoothed regime probabilities, $\Pr(s_T|\mathbf{Y}_{T+j})$ with $j > 0$. In other words, we can interpret the new datings as a '*real-time*' classification of the euro-zone business cycle, which only uses the information up to turning point.[3]

To gauge the reliability of the filtered regime probabilities in identifying new turning points, the resulting regime classification could be compared to the business cycle chronologies derived from the smoothed regime probabilities, which use the full-sample information.

3.5 Predicting the Business Cycle

Detecting recent regime shifts is essential to predict MS-VAR processes. A useful framework for the prediction of future regimes is the VAR(1) representation of the hidden Markov chain, where the discrete state variable s_t is substituted by the unobservable state vector ξ_t consisting of M binary indicator variables:

$$\xi_t = \left[\begin{array}{ccc} \mathrm{I}(s_t = 1) & \cdots & \mathrm{I}(s_t = M) \end{array} \right]'.$$

Let ξ_t be an $(M \times 1)$ vector whose ith element is unity when $s_t = i$, and zero otherwise. Then, the hidden Markov chain can be written as:

$$\xi_{t+1} = \mathbf{F}\xi_t + v_{t+1}, \tag{15}$$

where v_{t+1} is a martingale difference sequence and $\mathbf{F} = \mathbf{P}'$ is called the *'transition matrix'*. This representation allows the application of the well-established theory of forecasting linear systems to the problem of calculating future regime probabilities (see Krolzig (1997b)). If the Markov chain is ergodic, *i.e.* there does not exist an absorbing state, the process in (15) is stationary and will revert to its mean $\mathsf{E}[\xi_t] = \bar{\xi}$, where $\bar{\xi}$ is the vector of ergodic regime probabilities.

Since s_t is unobservable, so is ξ_t. Therefore we have to utilize the filter in (7) again, and work with the vector of filtered regime probabilities $\hat{\xi}_{t|t}$ instead:

$$\hat{\xi}_{t|t} = \left[\begin{array}{ccc} \Pr(s_t = 1|\mathbf{Y}_t) & \cdots & \Pr(s_t = M|\mathbf{Y}_t) \end{array} \right]'.$$

By using the linearity of the transition equation (15), the h-step prediction of ξ_{t+h} follows as:

$$\hat{\xi}_{t+h|t} = \mathbf{F}^h \hat{\xi}_{t|t}. \tag{16}$$

The predictive value of the regime inference, at the time when the prediction is made, depends positively on the persistence of the regime generating process. Under ergodicity, the regime probabilities converge to the unconditional ones:

$$\lim_{h \longrightarrow \infty} \mathbf{F}^h \hat{\xi}_{t|t} = \bar{\xi}.$$

[3]We still employ the full-sample parameter estimates.

For a two regime model, these ergodic probabilities are given by:

$$\bar{\xi}_1 = \frac{p_{21}}{p_{12+}p_{21}} \text{ and } \bar{\xi}_2 = \frac{p_{12}}{p_{12+}p_{21}}.$$

While the future distribution of the regimes is well defined, the model is less suited to forecast future turning points. In two regime model of § 2.1, '*peaks*' (transitions from regime 2 to 1) and '*troughs*' (transitions from regime 1 to 2) will on average occur with same probability:

$$\Pr(s_t = 1, s_{t+1} = 2) = \bar{\xi}_1 p_{12} = \frac{p_{21}p_{12}}{p_{12+}p_{21}} = \bar{\xi}_2 p_{21} = \Pr(s_t = 2, s_{t+1} = 1).$$

A related question is when the next turning point is likely to occur. Therefore it is useful to know the average duration h of '*expansions*' and 'contractions', which is given by:

$$\mathsf{E}[h_1] = \frac{1}{1 - p_{12}} \text{ and } \mathsf{E}[h_2] = \frac{1}{1 - p_{21}}.$$

Due to the stationary VAR(1) representation of an ergodic Markov chain, MS-VAR processes have short memory. The longer the forecast horizon, the better the linear approximation of the optimal predictor and the greater the potential unpredictability of (detrended) economic time series (see Krolzig (2003b), for further details on predicting MS-VAR processes). Finally, it is worth noting that the MS-VAR's ability of forecasting business cycle transitions can be enhanced further by making the transition probabilities time-varying (see, *inter alia*, Diebold, Lee, and Weinbach (1994)).

4 Properties of the Regime Classification Rule

In the following, we analyze the properties of the regime classification rule derived in the previous section. We are particularly interested in the dependence of the regime classification on (i) our beliefs regarding the state of the economy in the previous period and (ii) the current realization of the variables of interest. For sake of simplicity, we restrict our analysis to the '*real-time*' classification problem (filtering) and ignore the effects of future observations. We also abstract from estimation uncertainty by analyzing the situation when the true data generating process is an MS-VAR and the parameters are known. First, we consider a univariate Markov-switching model, and then the system case.

4.1 Univariate Markov-Switching Models

4.1.1 Analytics

The model we analyze here is kept general by defining it in terms of the conditional density of the variable of interest, y_t, in the two regimes, which in the context of business cycle analysis are supposed to represent recessions and expansions.

Regime 1 *(Recession)*:

$$\mathsf{p}\left(y_t \mid \mathbf{Y}_{t-1}, s_t = 1\right) = \frac{1}{\sigma_1}\phi\left(\frac{y_t - \mu_{1t}}{\sigma_1}\right) = f_1(y_t \mid \mathbf{Y}_{t-1}).$$

Regime 2 *(Expansion)*:

$$\mathsf{p}\left(y_t \mid \mathbf{Y}_{t-1}, s_t = 2\right) = \frac{1}{\sigma_2}\phi\left(\frac{y_t - \mu_{2t}}{\sigma_2}\right) = f_2(y_t \mid \mathbf{Y}_{t-1}).$$

where $\mu_{2t} > \mu_{1t}$ with μ_{2t} being positive, and μ_{1t} being negative if regime 1 is supposed to represent contractions in economic activity rather than low growth episodes; $\phi(\cdot)$ is the standard normal pdf.

As discussed in § 3.2, the observation of y_t will be classified as belonging to the regime 1, iff:

$$\frac{\Pr(s_t = 1 \mid \mathbf{Y}_t)}{\Pr(s_t = 2 \mid \mathbf{Y}_t)} = \frac{f_1(y_t|\mathbf{Y}_{t-1})\Pr(s_t = 1 \mid \mathbf{Y}_{t-1})}{f_2(y_t|\mathbf{Y}_{t-1})\Pr(s_t = 2 \mid \mathbf{Y}_{t-1})} > 1.$$

Under the assumed normality of $f_1(y_t \mid \mathbf{Y}_{t-1})$ and $f_2(y_t \mid \mathbf{Y}_{t-1})$, it can be shown that a linear discrimination rule emerges when the two regimes are homoscedastic, $\sigma_1^2 = \sigma_2^2 = \sigma^2$:

$$y_t < \frac{1}{2}\left(\mu_{2t} + \mu_{1t}\right) + \frac{\sigma^2}{\mu_{2t} - \mu_{1t}} \ln \pi, \tag{17}$$

where $\pi = (\Pr(s_t = 1|\mathbf{Y}_{t-1})/[1 - \Pr(s_t = 1|\mathbf{Y}_{t-1})])$ is the logit of the predicted probability of a recession formed at time $t-1$. Whenever y_t is less than the specified threshold, y_t is classified as belonging to the regime 1. The lower y_t, the higher the chance that the economy is in the state of a recession. The threshold depends on the parameters of the model and the predicted regime probabilities.

Suppose now that the two regimes are heteroscedastic: $\sigma_1^2 \neq \sigma_2^2$. Then, the condition for a recession is quadratic in y_t:

$$(\omega - 1)\, y_t^2 - 2\left(\omega\mu_{2t} - \mu_{1t}\right) y_t \quad > \quad -\left(\omega\mu_{2t}^2 - \mu_{1t}^2\right) + \sigma_1^2\left(\ln\omega - 2\ln\pi\right), \tag{18}$$

where $\omega = \sigma_1^2/\sigma_2^2$. If $\omega < 1$, the recessions are declared if y_t is in the interior region (y_t^-, y_t^+) with the threshold values:

$$y_t^{\pm} = \frac{\omega\mu_2 - \mu_1}{\omega - 1} \pm \frac{\sqrt{\omega}}{\omega - 1}\sqrt{(\omega - 1)\sigma^2\ln\omega - 2(\omega - 1)\sigma^2\ln\pi + (\mu_2 - \mu_1)^2}$$

Otherwise, recessions are found in the outer regions of $\mathbb{R}$. This highlights the danger of making both the first and second conditional moment of y_t regime dependent. Great care is required to ensure that the regime classification is economically meaningful (see Example 4 for a bivariate illustration).

Example 1: Hamilton Model

We evaluate the properties of the regime classification procedure in a simplified variant of Hamilton's business cycle model discussed in § 2.1:

Regime 1 *(Recession)*:

$$(y_t \mid \mathbf{Y}_{t-1}, s_t = 1) \sim \mathsf{N}(-0.4, 0.5^2).$$

Regime 2 (*Expansion*):

$$(y_t \mid \mathbf{Y}_{t-1}, s_t = 2) \sim \mathsf{N}(+1.2, 0.5^2).$$

Markov chain (*Transition probabilities*):

$$\begin{aligned} p_{11} = \Pr(s_t = 1 \mid s_{t-1} = 1) &= 0.7, \\ p_{22} = \Pr(s_t = 2 \mid s_{t-1} = 2) &= 0.9. \end{aligned}$$

Expansions are more persistent than recessions. For the transition probabilities given, the ergodic probability of a recession is 0.25.

The mechanics of the result in (17) are illustrated in Figure 1. The first panel shows the conditional density of the variable of interest, y_t, in recessions ($s_t = 1$) and in expansions ($s_t = 2$). Suppose now that the regime probabilities at time $t-1$ are known. Then the predicted regime probabilities, $\Pr(s_t|Y_{t-1})$, result as follows:

$$\begin{aligned} \Pr(s_t = 1 \mid \mathbf{Y}_{t-1}) &= 0.7\Pr(s_{t-1} = 1 \mid \mathbf{Y}_{t-1}) + 0.1\Pr(s_{t-1} = 2 \mid \mathbf{Y}_{t-1}), \\ \Pr(s_t = 2 \mid \mathbf{Y}_{t-1}) &= 0.3\Pr(s_{t-1} = 1 \mid \mathbf{Y}_{t-1}) + 0.9\Pr(s_{t-1} = 2 \mid \mathbf{Y}_{t-1}), \end{aligned}$$

where we used that $\Pr(s_t = j \mid \mathbf{Y}_{t-1}) = \sum_{i=1}^{M} p_{ij} \Pr(s_{t-1} = i \mid \mathbf{Y}_{t-1})$.

The predicted density of y_t is given by equation (8) as:

$$\mathsf{p}(y_t \mid \mathbf{Y}_{t-1}) = \sum_{j=1}^{M} \Pr(s_t = j \mid \mathbf{Y}_{t-1}) f_j(y_t|\mathbf{Y}_{t-1}),$$

which is a mixture of normals. In the second panel of Figure 1, the predicted density is plotted for different values of the filtered probability of a recession in the previous period, $\Pr(s_{t-1} = 1 \mid \mathbf{Y}_{t-1}) = w_{1t}$. Those determine the weights of the two Gaussian densities of y_t conditioned on s_t:

$$p(y_t \mid \mathbf{Y}_{t-1}) = [0.1 + 0.6w_{1t}] f_1(y_t|\mathbf{Y}_{t-1}) + [0.9 - 0.6w_{1t}] f_2(y_t|\mathbf{Y}_{t-1}).$$

The third panel shows the resulting filtered probability of a recession ($s_t = 1$) given the observed value of y_t according to (7). Finally notice how the classification rule proposed in (14) implies a unique threshold y^* such that for $y_t < y^*$ ($y_t \geq y^*$), the econometrician would conclude that the economy is in the state of a recession (expansion).

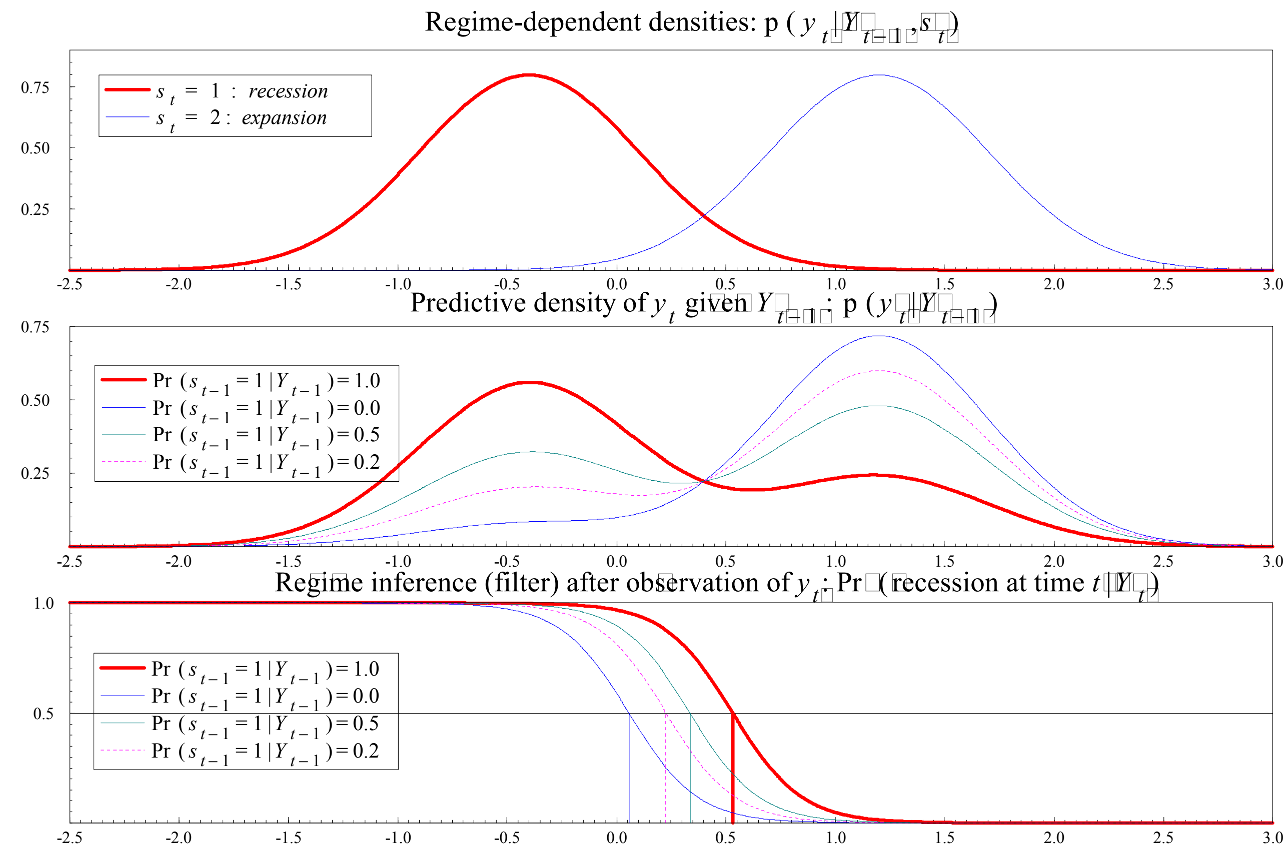

Figure 1: Regime inference in univariate Markov-switching models

4.2 Regime Classification in an MS-VAR Model

4.2.1 Analytics

We now extend the analysis of the regime classification in Markov-switching models, applied so far to univariate models, to the vector case. Consider the following two-regime MS-VAR with:

Regime 1 *(Recession)*:

$$(\mathbf{y}_t | \, \mathbf{Y}_{t-1}, s_t = 1) \sim \mathsf{N}(\boldsymbol{\mu}_{1t}, \boldsymbol{\Sigma}_1);$$

Regime 2 (*Expansion*):

$$(\mathbf{y}_t | \, \mathbf{Y}_{t-1}, s_t = 2) \sim \mathsf{N}(\boldsymbol{\mu}_{2t}, \boldsymbol{\Sigma}_2),$$

where $\boldsymbol{\mu}_{2t} > \boldsymbol{\mu}_{1t}$ is the conditional mean of $\mathbf{y}_t$.

As in § 4.1.1, the observation of $\mathbf{y}_t = (y_{1t}, \ldots, y_{Kt})'$ will be classified as belonging to the regime 1 iff the likelihood of regime 1 is greater than the likelihood of regime 2:

$$\frac{\Pr(s_t = 1 \mid \mathbf{Y}_t)}{\Pr(s_t = 2 \mid \mathbf{Y}_t)} = \frac{f_1(\mathbf{y}_t | \mathbf{Y}_{t-1}) \Pr(s_t = 1 \mid \mathbf{Y}_{t-1})}{f_2(\mathbf{y}_t | \mathbf{Y}_{t-1}) \Pr(s_t = 2 \mid \mathbf{Y}_{t-1})} > 1.$$

Under normality of $f_1(\mathbf{y}_t|\mathbf{Y}_{t-1})$ and $f_2(\mathbf{y}_t|\mathbf{Y}_{t-1})$, parametric conditions can be derived for the values of $\mathbf{y}_t$ leading to the conclusion that the economy is stuck in a recession. If the two regimes are homoscedastic, *i.e.* $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$, it can be shown that again a linear discrimination rule results (see Krolzig (2003a), for details):

$$(\boldsymbol{\mu}_{2t} - \boldsymbol{\mu}_{1t})' \mathbf{\Sigma}^{-1}\mathbf{y}_t < \frac{1}{2}\left(\boldsymbol{\mu}_{2t}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_{2t} - \boldsymbol{\mu}_{1t}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_{1t}\right) + \ln\pi, \tag{19}$$

where $\pi = \Pr(s_t = 1|\mathbf{Y}_{t-1})/\left[1 - \Pr(s_t = 1|\mathbf{Y}_{t-1})\right]$. For bivariate processes, $K = 2$, with a diagonal variance-covariance matrix,*i.e.* $\sigma_{kl} = 0$ for $k \neq l$, the rule can be further simplified to:

$$\begin{aligned}
\mathbf{b}_t'\mathbf{y}_t &< c_t, \qquad (20)\\
\text{with:}\quad \mathbf{b}_t &= \left[\ \sigma_1^{-2}\left(\mu_{1.2t} - \mu_{1.1t}\right), \sigma_2^{-2}\left(\mu_{2.2t} - \mu_{2.1t}\right)\ \right]',\\
c_t &= \frac{1}{2}\left(\frac{\mu_{1.2t}^2 - \mu_{1.1t}^2}{\sigma_1^2} + \frac{\mu_{2.2t}^2 - \mu_{2.1t}^2}{\sigma_2^2}\right) + \ln\left(\frac{\Pr(s_t = 1|\mathbf{Y}_{t-1})}{1 - \Pr(s_t = 1|\mathbf{Y}_{t-1})}\right).
\end{aligned}$$

The linearity of the divide between recessions and expansions can be easily spotted in the examples given in Figures 4 and 7, which will be discussed next. For a heteroscedastic error process, $\mathbf{\Sigma}_1 \neq \mathbf{\Sigma}_2$, the condition for a recession is quadratic in $\mathbf{y}_t$; this will be investigated further in the fourth example (see, also, Figure 10).

It is worth remembering that the regime classification rule (19) for MS-VAR models is independent of the weight of any country in the index. Indeed, scaling one of the countries would result in exactly the same regime classification. However, note that larger countries tend to be associated with less noisy time series than smaller countries. Fluctuations in the former countries will deliver clearer signals and, therefore, will usually have a stronger impact on the regime probabilities. Furthermore, the stronger the country is synchronized with the common business cycle and the more pronounced the cycle affects the country's economy; the higher is the weight of the country in the discriminator function.

This has serious implications for business cycle analysis with aggregated data (see Krolzig (2003a)). The question is whether the aggregation of the country data $\mathbf{y}_t = (y_{1t}, \ldots, y_{Kt})'$ to an index y_t^a affects the efficiency of the detection of business cycle turning points, when $\mathbf{y}_t$ is generated by an MS-VAR as defined in (6). In other words, can the aggregate y_t^a be a sufficient statistic for $\hat{s}_t$? It can be shown that the answer to this question depends on:

$$\Pr\left(s_t|y_t^a\right) = \Pr\left(s_t|\mathbf{Y}_t\right), \tag{21}$$

which is investigated in Krolzig (2003a). In the case of a heteroscedastic error term, any linear aggregation of the data $y_t^a = \mathbf{a}'\mathbf{y}_t$ must clearly be inefficient.

4.2.2 Example 2: Bivariate Markov-Switching Model

The regime classification in MS-VAR models is illustrated by three variations of a bivariate MS-VAR process, which generalizes the univariate business cycle model considered earlier. We can think here of two countries driven by the same business cycle, such that y_{kt} denotes the growth rate of output in country $k \in \{1, 2\}$. We assume that the distribution of $\mathbf{y}_t = (y_{1t}, y_{2t})'$

is Gaussian in both regimes, and that output contracts by 0.4 percent per quarter in *recessions* and grows by 1.2 percent in *expansions*. The standard deviation of output growth is 0.5 percent in both regimes.

Regime 1 *(Recession)*:

$$(\mathbf{y}_t \mid s_t = 1) \sim \mathsf{N}\left(\begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix}, \begin{bmatrix} 0.5^2 & 0 \\ 0 & 0.5^2 \end{bmatrix} \right).$$

Regime 2 (*Expansion*):

$$(\mathbf{y}_t \mid s_t = 2) \sim \mathsf{N}\left(\begin{bmatrix} 1.2 \\ 1.2 \end{bmatrix}, \begin{bmatrix} 0.5^2 & 0 \\ 0 & 0.5^2 \end{bmatrix} \right).$$

The hidden *Markov chain* generating the regimes is chosen to reflect the properties of the US business cycle as found in Hamilton (1989):

$$\begin{aligned} p_{11} = \Pr(s_t = 1 \mid s_{t-1} = 1) &= 0.7, \\ p_{22} = \Pr(s_t = 2 \mid s_{t-1} = 2) &= 0.9. \end{aligned}$$

Figure 2 plots the probability density function (pdf) of $\mathbf{y}_t$ in *recessions*, $s_t = 1 : f_1(\mathbf{y}_t|\mathbf{Y}_{t-1})$,and in *expansions*, $s_t = 2 : f_2(\mathbf{y}_t|\mathbf{Y}_{t-1})$. Given the state of the business cycle, the density of y_{1t} and y_{2t} is simply Gaussian with mean vector $\boldsymbol{\mu}_1$ or $\boldsymbol{\mu}_2$, and variance matrix $\mathbf{\Sigma}$.

Since the regime itself is unobservable, our knowledge of the distribution of $\mathbf{y}_t$ given the information available at time $t-1$ is in fact much more limited. The predicted density of $\mathbf{y}_t$ plotted in Figure 3 is a mixture of normals:

$$\mathsf{p}(\mathbf{y}_t \mid \mathbf{Y}_{t-1}) = w_{1t} f_1(y_t|\mathbf{Y}_{t-1}) + (1 - w_{1t}) f_2(y_t|\mathbf{Y}_{t-1}),$$

where w_{1t} denotes the weight attached to conditional pdf of regime 1 at time t.

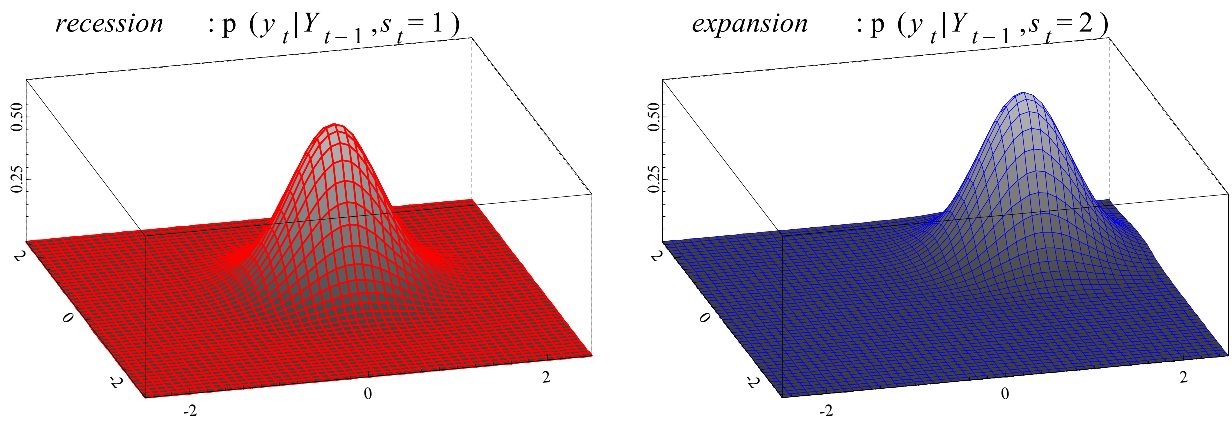

Figure 2: Probability density of y_t in recessions and expansions

The weight w_{1t} of regime 1 is given by its predicted regime probability, $\Pr(s_t = 1 \mid \mathbf{Y}_{t-1})$, which is determined by the regime inference at time and the structure of the matrix of transition probabilities:

$$\begin{aligned}\Pr(s_t = 1 \mid \mathbf{Y}_{t-1}) &= p_{11} \Pr(s_{t-1} = 1 \mid \mathbf{Y}_{t-1}) + p_{21} \Pr(s_{t-1} = 2 \mid \mathbf{Y}_{t-1}) \\ &= p_{21} + (p_{11} - p_{21}) \Pr(s_{t-1} = 1 \mid \mathbf{Y}_{t-1}).\end{aligned}$$

Using (7), we can use this information to calculate the filtered probability of the regimes – shown in Figure 4 for regime 2:

$$\Pr(s_t = 2|\mathbf{y}_t, \mathbf{Y}_{t-1}) = \frac{(1 - w_{1t})\ f_2(\mathbf{y}_t|\mathbf{Y}_{t-1})}{w_{1t}\ f_1(\mathbf{y}_t|\mathbf{Y}_{t-1}) + (1 - w_{1t})\ f_2(\mathbf{y}_t|\mathbf{Y}_{t-1})}.$$

The $\Pr(s_t = 2|\mathbf{y}_t, \mathbf{Y}_{t-1}) = 0.5$ isolate, which gives the classification into recessions and contractions, is a linear function of y_{1t} and y_{2t}.

4.2.3 Example 3: Lack of Business Cycle Synchronization

Here, we alternate Example 2 by limiting the business cycle to the first variable:

$$\boldsymbol{\mu}_1 = \begin{bmatrix} -0.4 \\ 1.2 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 1.2 \\ 1.2 \end{bmatrix}.$$

The same calculations as in the previous example apply. Since the mean growth rate of the second variable is unaffected by shifts in the regime variable s_t, we have that the unconditional pdf of y_{1t} and y_{2t} shown in Figure 5, just shifts in the horizontal direction whenever there is a transition in the state of the business cycle. The resulting predicted density depicted in Figure 6 is clearly seen to be bimodal with both modes lying on a horizontal line.

Figure 7 again plots the filtered probability of an *expansion* (regime 2) at time t conditional on the probability of a *recession* (regime 1) at time $t-1$ and the realizations of y_{1t} and y_{2t}. We are again interested in what we can learn about the state of business cycle from the observed values of y_{1t} and y_{2t}. It turns out that y_{2t} does not contribute to the understanding of the underlying state of the economy. The reason for this is that since y_{2t} does not depend on s_t, observing y_{2t} can not reveal any information about the current realization of s_t:

$$\Pr(s_t|y_{1t}, y_{2t,}\mathbf{Y}_{t-1}) = \Pr(s_t|y_{1t}, \mathbf{Y}_{t-1}).$$

Hence, modelling y_{2t} does not contribute to the solution of the signal extraction problem. The regime inference just depends on y_{1t} in the same way as in the simplified Hamilton model considered in Example 1.

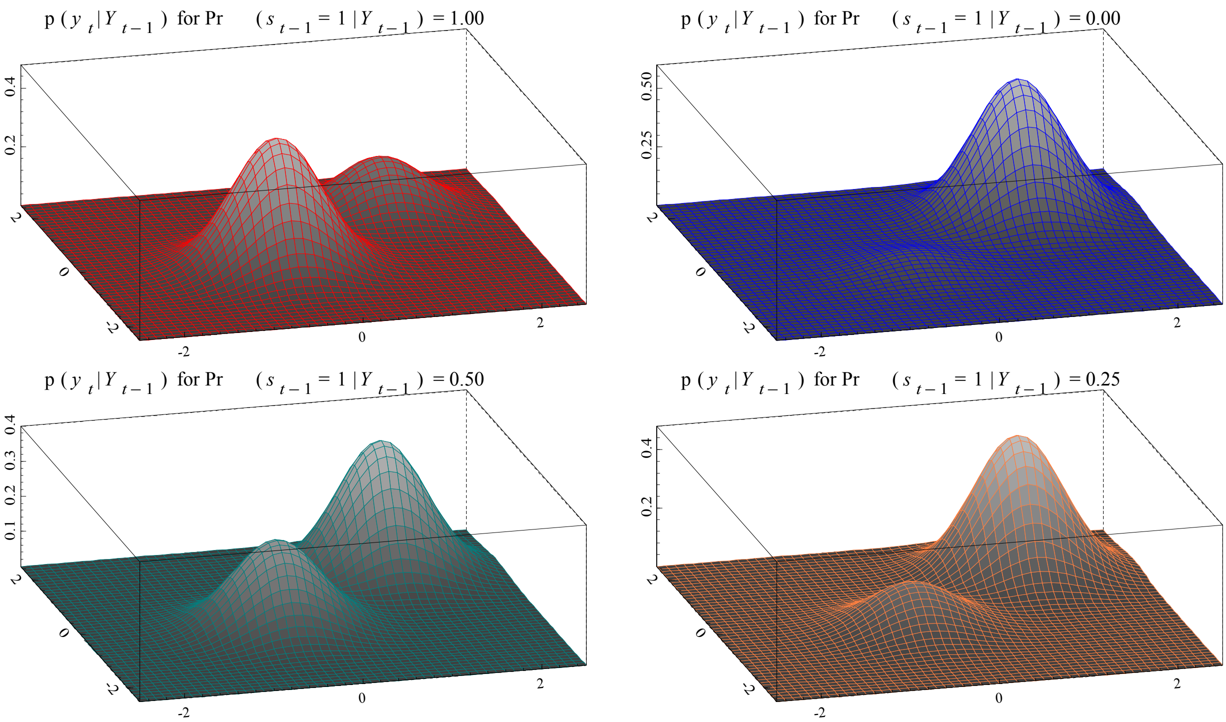

Figure 3: Predicted density of y_t given Y_{t-1} and the probability of recessions at time $t-1$

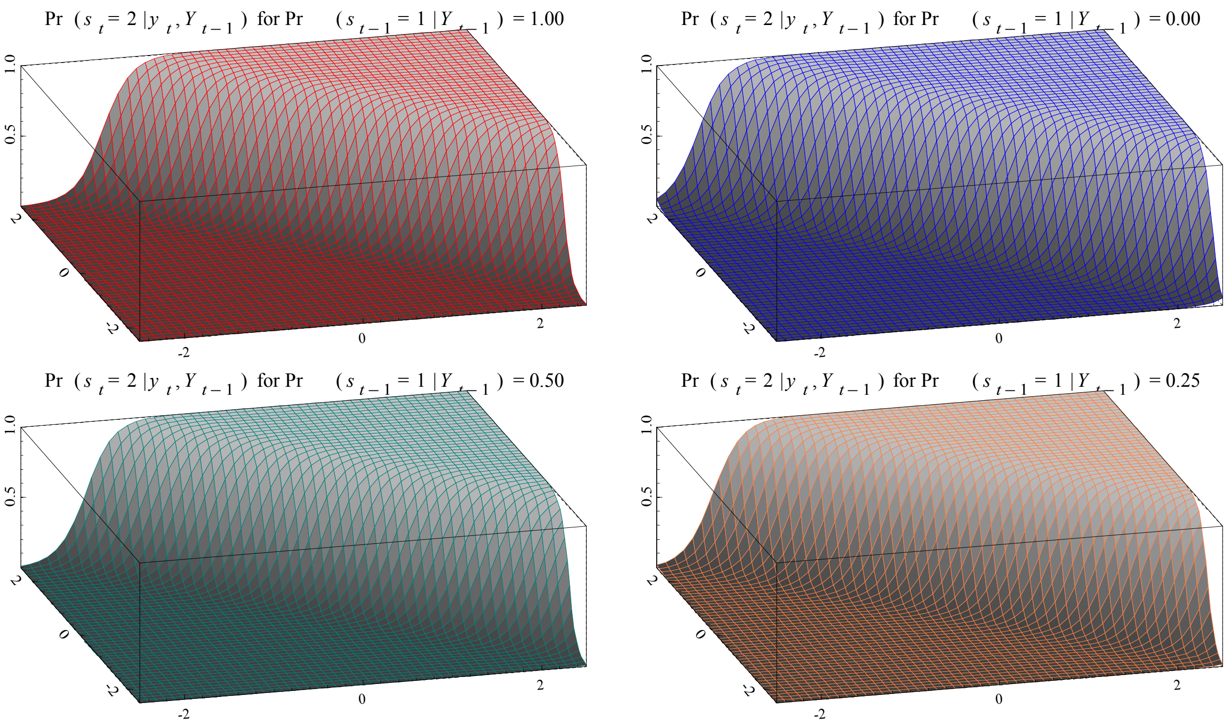

Figure 4: Filtered probability of regime 2 (expansion) inferred from the observation of y_t

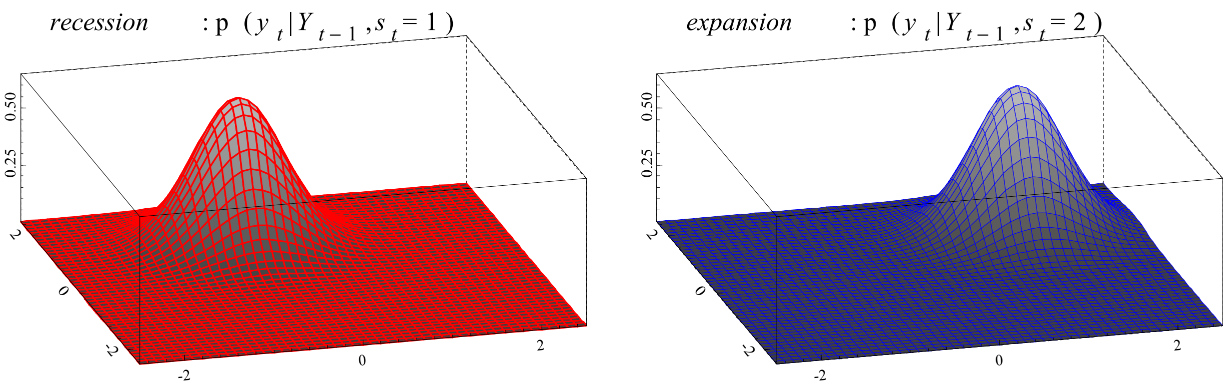

Figure 5: Probability density of y_t in recessions and expansions for $\mu_{1.1} = \mu_{1.2}$

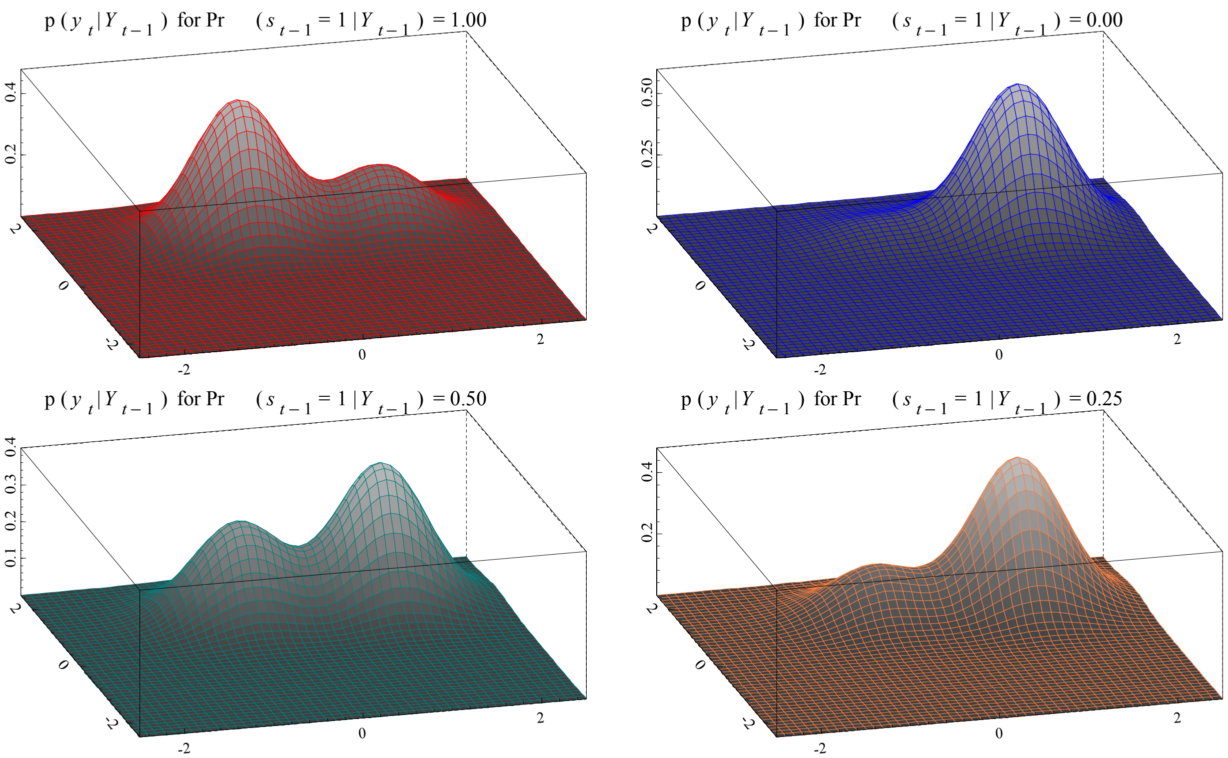

Figure 6: Predicted density of y_t given Y_{t-1} and the probability of recessions at time $t-1$ for $\mu_{1.1} = \mu_{1.2}$

4.2.4 Example 4: Regime-Dependent Heteroscedasticity

This final example highlights the potential dangers in interpreting Markov-switching models as representations of the business cycle when they are characterized by time-varying volatility (regime-dependent variances).[4]

The MS-VAR here differs from the model in the second example by allowing for regime-dependent heteroscedasticity:

$$\mathbf{\Sigma}_1 = \begin{bmatrix} 0.5^2 & 0 \\ 0 & 0.5^2 \end{bmatrix}, \quad \mathbf{\Sigma}_2 = \begin{bmatrix} 0.8^2 & 0 \\ 0 & 0.8^2 \end{bmatrix}.$$

So the expansionary regime is characterized by a higher variance. The mean vectors are unchanged:

$$\boldsymbol{\mu}_1 = \begin{bmatrix} -0.4 \\ 1.2 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} -0.4 \\ 1.2 \end{bmatrix}.$$

The conditional density of y_{1t} and y_{2t} in the two regimes is depicted in Figure 8.

If the regime were known, the conditional variance of y_t would be higher in recessions than in expansions. This is, however, not so trivial when we have to rely on statistical inference regarding the unobserved regime variable.

The predicted density of $\mathbf{y}_t$ given its history and the probability of regime 1 at time $t-1$ is shown in Figure 9. The pdf is bimodal, with the weight of the two conditional densities being determined by $w_{1t} = \Pr(s_t = 1 \mid Y_{t-1})$. So the conditional variance of y_{1t} and y_{2t} is given by:

$$\begin{aligned} \mathsf{Var}\left[y_{kt}|\mathbf{Y}_{t-1}\right] &= \mathsf{E}_{s_t}\left\{\mathsf{Var}\left[y_{kt}|s_t,\mathbf{Y}_{t-1}\right]\right\} + \mathsf{Var}_{s_t}\left\{\mathsf{E}\left[y_{kt}|s_t,\mathbf{Y}_{t-1}\right]\right\} \\ &= w_{1t}\sigma_1^2 + (1-w_{1t})\sigma_2^2 + w_{1t}(\mu_1 - \mathsf{E}\left[y_{kt}|\mathbf{Y}_{t-1}\right])^2 + (1-w_{1t})(\mu_2 - \mathsf{E}\left[y_{kt}|\mathbf{Y}_{t-1}\right])^2 \\ &= \sigma_2^2 + (\sigma_1^2 - \sigma_2^2)w_{1t} + (\mu_1 - \mu_2)^2(1-w_{1t})w_{1t} \\ &= 0.64 - 2.56w_{1t} + 2.17w_{1t}^2. \end{aligned}$$

It is highest for $\Pr(s_t = 1 \mid \mathbf{Y}_{t-1}) \approx 0.424$, or in other words when the available information at time $t-1$ is inconclusive regarding the current state of the economy: $\Pr(s_{t-1} = 1 \mid \mathbf{Y}_{t-1}) \approx 0.54$. This is expressed in Figure 9 in the widespread distribution of probability mass for $\Pr(s_{t-1} = 1 \mid \mathbf{Y}_{t-1}) = 0.5$.

[4] Similar problems can result from making the VAR coefficients regime dependent.

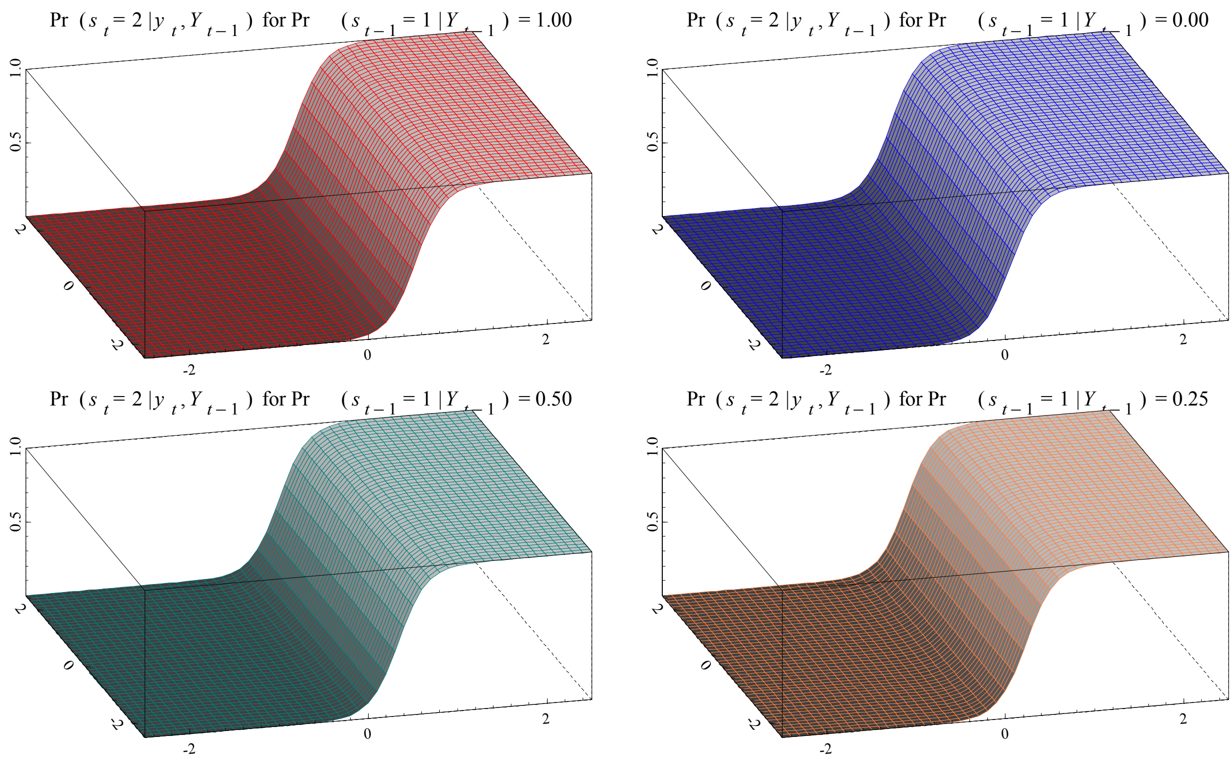

Figure 7: Filtered probability of regime 2 (expansion) inferred from the observation of y_t for $\mu_{1.1}= \mu_{1.2}$

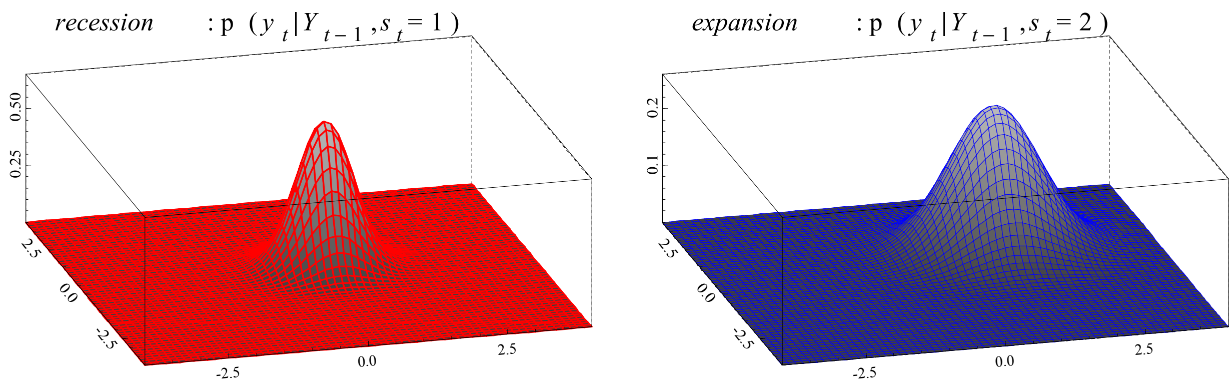

Figure 8: Probability density of y_t in recessions and expansions for $\Sigma_2 > \Sigma_1$

The signal extraction problem has become more complicated when compared to the previous examples. Due to the regime-dependent heteroscedasticity, regime 1 exhibits not only lower mean growth rates than regime 2, but also a lower variance. In general, negative growth rates signal *recessions* (regime 1). However, if the observed growth rates of y_{1t} and y_{2t} become increasingly low (*i.e.* negative), the filter reinterprets the situation: while $\mathbf{y}_t$ appears to be an outlier from the perspective of regime 1, it is still quite likely under regime 2 due to its greater variance. This in turn makes regime 2 look more likely to prevail. The outcome of the filter is recorded in Figure 10, which displays the probability of regime 2. It is worth remembering that great care is needed when interpreting the business cycle features of an MS-VAR model with switching variance or switching coefficients.

4.3 Summary

The results derived here have decisive implications for the detection of business cycle turning points:

i. Importance of parsimony and meaningfulness of the regimes.

ii. General inefficiency of aggregation.

iii. The stronger the country is synchronized with the common business cycle and the more pronounced the cycle affects the country's economy, the higher is the weight of the country in the regime classification.

iv. Regime classification is independent of the size of the countries or their weight in an index or aggregate (such as euro-zone GDP).

The importance of these issues will become evident in the following empirical analysis of the euro-zone business cycle.

5 The Euro-Zone Business Cycle

Throughout in this paper, we emphasized the importance of modelling systems for the construction of turning point chronologies of the euro-zone business cycle. In § 2 we argued that the euro-zone business cycle constitutes a common feature of the macroeconomic dynamics in the EMU member states. In contrast to the analyses of aggregated euro-zone data, this notion is reflected in Markov-switching vector autoregressive models of macroeconomic activity in euro-zone countries. The results derived in § 4 provided further reasons for analyzing systems instead of aggregates, foremost the missing invariance of the regime classification under aggregation. In this section we will first discuss various empirical Markov-switching models of the European (EU/EMU) business cycle proposed in the literature. We then present the empirical modelling approach, which will be applied to three different statistical measures of macroeconomic activity in EMU member states. Finally, we present the empirical results of this paper derived from the estimated Markov-switching vector autoregressive models of the euro-zone business cycle.

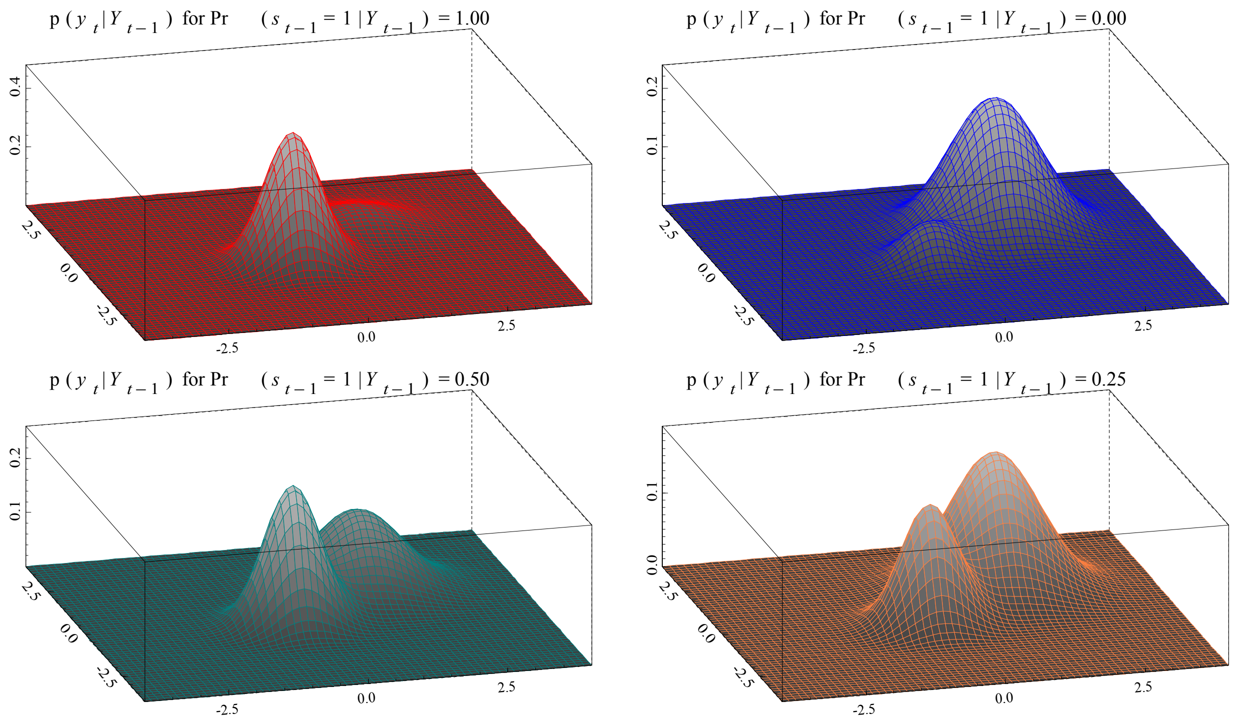

Figure 9: Predicted density of y_t given Y_{t-1} and the probability of recessions at time $t-1$ for $\Sigma_2 > \Sigma_1$

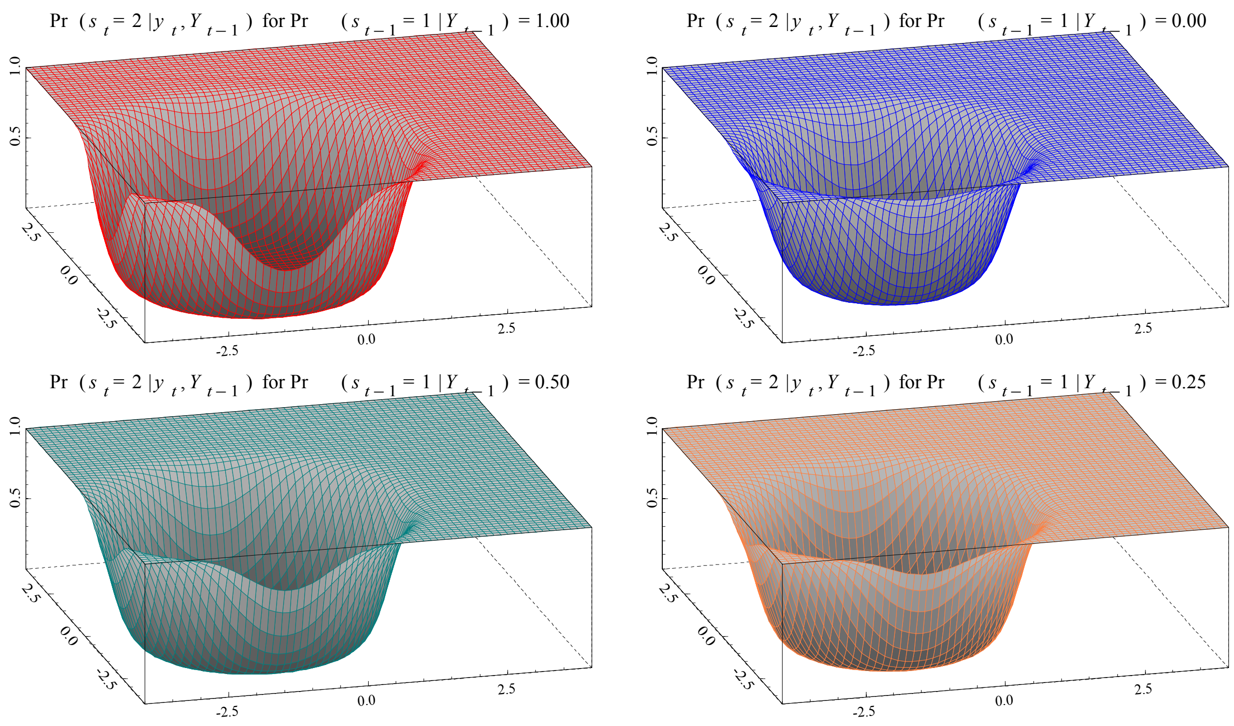

Figure 10: Filtered probability of regime 2 (expansion) inferred from the observation of y_t for $\Sigma_2 > \Sigma_1$

5.1 Markov-Switching Studies of the European Business Cycle

Recent years have seen an increasing attempt to investigate international business cycles with modern non-linear time series models. First contributions to the analysis of international business cycles with Markov-switching models were made by Phillips (1991) and Filardo and Gordon (1994). With the Treaty of Maastricht, the focus shifted towards the analysis of business cycle synchronization in the European Union, and later in the euro zone. Table 1 gives an overview on the recent literature on Markov-switching models of the European Business Cycle.

Artis, Krolzig, and Toro (2004) investigated the existence and identification of a European Business Cycle by analyzing monthly industrial production data of nine EU countries from 1970M01 to 1996M12.

Two important issues arose in their investigation: (i) the convergence process of Southern Europe and (ii) the secular decline of the mean growth rates in the post-Bretton Woods era. Analyzing GDP data for six EU countries from 1970Q3 to 1995Q4, Krolzig (2001a) concludes that two-regime models representing contractions and expansions are inconsistent with these two stylized facts of the post-war economic history of Western Europe. Krolzig and Toro (2000) compare the 'classical' approach proposed in Burns and Mitchell (1946) of dating and analyzing the business cycle with the Markov-switching approach as its 'modern' alternative.

By using the model's regime probabilities as an optimal statistical inference of the turning point of the European business cycle, they demonstrate the capacity of the MS-VAR approach to generate the stylized facts of the *classical* cycle in Europe. These studies have in common the presence of a third regime, which however is rarely found after 1980.

We can therefore presume that for data beyond that period two-regime models provide an adequate description of the business cycle. Consequently, Krolzig (2001b) have fitted two-regime Markov-switching models to aggregated euro-zone real GDP data and disaggregated single-country euro-zone for the 1980Q2 to 2000Q4 period. Recessions in the euro-zone are found from 1980Q1 to 1981Q1 and 1992Q3 to 1993Q2. Also, Peersman and Smets (2001) find that a two-regime MS-VAR model delivers reasonable results for quarter-to-quarter growth rates of detrended industrial production in seven euro-area countries of the period 1978Q4 – 1998Q4.

In Figure 11, the (smoothed) probability of a recession in the studies of Artis, Krolzig, and Toro (2004), Krolzig and Toro (2000), and Krolzig (2001b) is plotted. When taking account of the different sample periods involved, the regime classifications consistently support the presence of three major recessions in the periods 1973/74, 1980-82 and 1991 to 1994, though there is increased uncertainty regarding the recession in the 1990s. Anas and Ferrara (2002) extend the univariate analysis of Krolzig (2001b) up to 2002Q2.

Using aggregated euro-zone GDP data, they discover a more recent recession starting in 2001Q1, with somehow different results when analyzing euro-zone IIP data. In the following, we reconsider the issue by analyzing three new multivariate data sets. An important question we will investigate is whether the turning point chronology remains robust when more recent economic data are considered.

Table 1: MS-Studies of the European business cycle

Model		Data	M	Sample	DE	FR	IT	ES	NL	BE	AT	FI	PT	GR	IE	LU	UK
MSIH(3)-VAR(4)	(22)	GDP[a]	3	'73.3-'02.1	+	+	+	+	+	−	+	−	−	−	−	−	−
MSM(2)–VAR(2)	(23)	IIP[a]	2	'74.1-'02.1	+	+	−	+	−	+	+	−	+	+	−	+	−
MSI(2)–VAR(2)	(24)	GDP[b]	2	'80.3-'02.1	+	+	+	+	+	+	−	+	−	−	−	−	−
Krolzig (2001b)	(K2)	GDP	2	'80.2-'00.4	+	+	+	+	+	+	+	+	−	−	−	−	−
Peersman and Smets (2001)	(PS)	IIP	2	'78.4-'98.4	+	+	+	+	+	+	+	−	−	−	−	−	−
Artis et al. (2004)	(AKT)	IIP	3	'70.1-'96.12	+	+	+	+	+	+	+	−	+	−	−	−	+
Krolzig and Toro (2000)	(KT)	GDP	3	'70.3-'95.4	+	+	+	+	−	−	+	−	−	−	−	−	+
Krolzig (2001a) Krolzig (2001b)	(K1)	GDP[c]	2	'80.2-'00.4													
Anas and Ferrara (2002)*	(AF1)	GDP[c]	2	'80.1-'02.2													
Anas and Ferrara (2002)	(AF2)	IIP[c]	3	'71.1-'02.8													

Notes: Anas and Ferrara (2002) also consider a three-regime model which is less robust. M: number of regimes. +: country is included. - country is excluded. IIP: industrial production index. GDP: gross domestic product (real). [a]: OECD. [b]: Eurostat.[c]: agg..

5.2 The Empirical Modelling Approach

In this paper, we follow the modelling approach taken in Krolzig (1997a), featuring a data-driven model specification as the basis for the derivation of economically meaningful results. For the dating of the European (EU) business cycle, Artis, Krolzig, and Toro (2004) proposed the following empirical modelling procedure:

i. *Data transformation*

- Seasonal adjustment (here: X-12-ARIMA);
- Outlier correction (here: *PcGets*);
- Smoothing (*e.g.*, centered MA of monthly data);
- Differencing.

ii. *Preliminary data analysis*

- Co-switching analysis (using univariate Markov-switching models);
- Co-integration analysis (using linear VAR models).

iii. *Markov-switching vector autoregressions*

- Maximum Likelihood estimation;
- Specification analysis.

iv. *Constructing the turning point chronology*

- Regime inference: filtering and smoothing;
- Business cycle chronology: regime classification and turning point dating;
- Control of the regime classification;
- Turning point detection.It is beyond the scope of this paper to discuss all the issues involved. The construction of the turning point chronology has been presented in the previous sections of this paper. A general overview and discussion of the econometric issues can be found in Krolzig (1997b). Readers, who might want to get some impression involved in practical use of these models, are referred to Artis, Krolzig, and Toro (2004).

Here we shall focus on the Maximum Likelihood (ML) estimation of the parameters of the MS-VAR model.[5] For the sample $\mathbf{Y}_T = (\mathbf{y}'_T, \ldots, \mathbf{y}'_1)'$ and initial values $(\mathbf{y}_0, \ldots, \mathbf{y}_{1-p})'$, the likelihood function is given by:

$$L(\boldsymbol{\lambda}) = \prod_{t=1}^{T} \sum_{i=1}^{M} f(\mathbf{y}_t \mid \mathbf{Y}_{t-1}; \boldsymbol{\theta}_m, \boldsymbol{\theta}_c) \Pr(s_t = i \mid \mathbf{Y}_{t-1}; \boldsymbol{\lambda}),$$

where $f(\mathbf{y}_t \mid \mathbf{Y}_{t-1}; \boldsymbol{\theta}_m, \boldsymbol{\theta}_c)$ is the pdf of $\mathbf{y}_t$ in regime m, $f_m(\mathbf{y}_t \mid \mathbf{Y}_{t-1})$. The parameter vector $\boldsymbol{\lambda} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M; \boldsymbol{\theta}_c; \{p_{ij}\}; s_0)'$ includes:

$\{p_{ij}\}$	set of free transition probabilities,
$\boldsymbol{\theta}_m$	vector of regime-dependent VAR parameters,
$\boldsymbol{\theta}_c$	vector of regime-invariant (common) parameters,
s_0	initial state of the regime variable.

Note that in MSM-VAR models (with switching means), the problem becomes more involved since $\mathbf{y}_t$ does not only depend on s_t, but also on $s_{t-1}, \ldots, s_{t-p}$. Due to the structure of the likelihood function, the first order conditions of the ML estimator are highly non-linear. So any estimation procedure is faced with the typical problems of non-linear optimization.

The Expectation-Maximization (EM) algorithm considered by Hamilton (1990) is a very robust technique to derive the ML estimates.

[5] We assume throughout that the data generating process is stationary, *i.e.* the Markov chain is ergodic and the roots of the characteristic VAR polynomial are outside the unit circle (see Krolzig (1996), for Markov-switching vector equilibrium correction models for cointegrated variables).

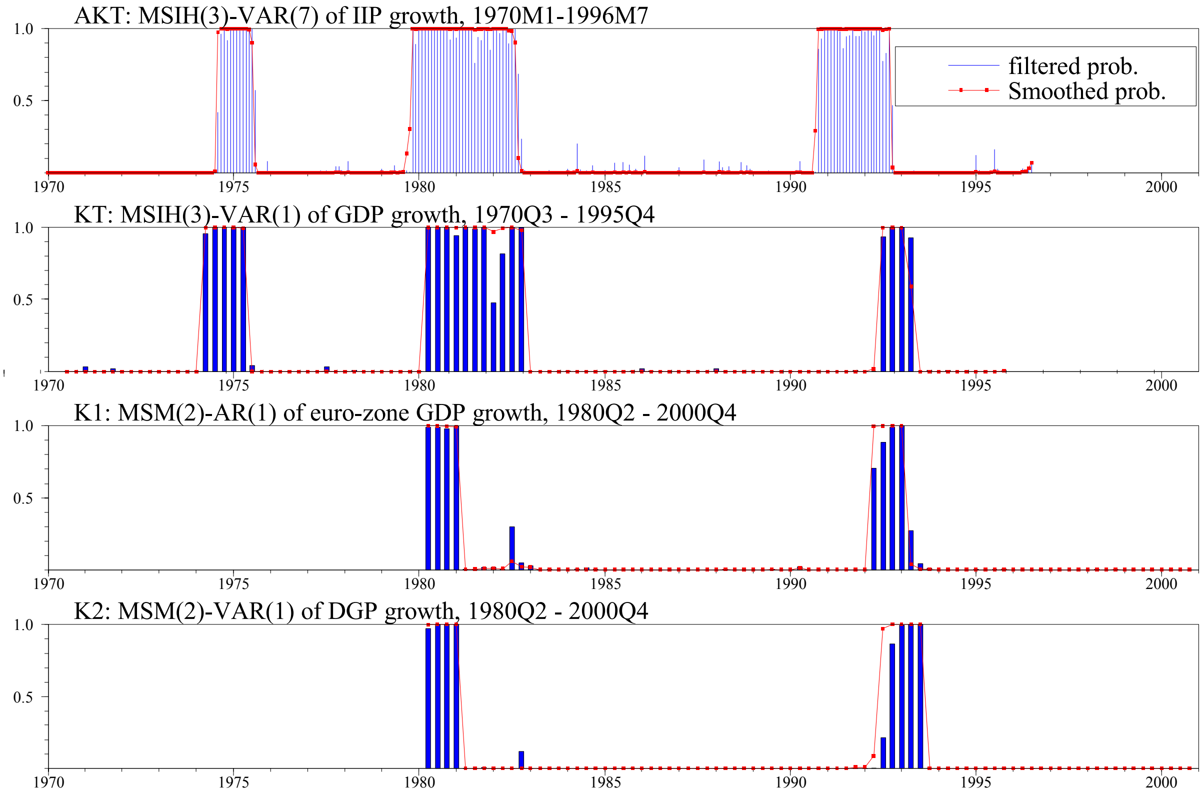

Figure 11: Classifications of the euro-zone business cycle in the literature: the probability of a recession

It consists of zigzag procedure iterating between an *Expectation step,* which runs the regime inference given the parameters of the previous iteration, and a *Maximization step*, which in turn uses the resulting smoothed regime probabilities to perform the parameter estimation:

i. Transition probabilities:

$$\tilde{p}_{ij}^{(n)} = \frac{\sum_{t=1}^{T-1} \Pr\left(s_t = i, s_{t+1} = j \mid \mathbf{Y}_T; \tilde{\boldsymbol{\lambda}}^{(n-1)}\right)}{\sum_{t=1}^{T-1} \Pr\left(s_t = i \mid \mathbf{Y}_T; \tilde{\boldsymbol{\lambda}}^{(n-1)}\right)}.$$

ii. VAR parameters:

$$\begin{aligned} \tilde{\boldsymbol{\theta}}_i^{(n)} &= \arg\max_{\boldsymbol{\theta}_i} \sum_{t=1}^{T} \ln f(\mathbf{y}_t \mid \mathbf{Y}_{t-1}; \boldsymbol{\theta}_m, \boldsymbol{\theta}_c) \Pr\left(s_t = i \mid \mathbf{Y}_T; \tilde{\boldsymbol{\lambda}}^{(n-1)}\right); \\ \tilde{\boldsymbol{\theta}}_c^{(n)} &= \arg\max_{\boldsymbol{\theta}_i} \sum_{t=1}^{T} \sum_{i=1}^{M} \ln f(\mathbf{y}_t \mid \mathbf{Y}_{t-1}; \boldsymbol{\theta}_m, \boldsymbol{\theta}_c) \Pr\left(s_t = i \mid \mathbf{Y}_T; \tilde{\boldsymbol{\lambda}}^{(n-1)}\right). \end{aligned}$$

iii. Initial State:

(a) ergodic regime probability: $\tilde{p}_{i,0}^{(n)} = \Pr\left(s_t = i \mid \boldsymbol{\lambda}^{(n-1)}\right)$;

(b) ML estimate: $\tilde{s}_0^{(n)} = \arg\max_{s_0 \in \{1,\ldots,M\}} L\left(\boldsymbol{\theta}_1^{(n-1)}, \ldots, \boldsymbol{\theta}_M^{(n-1)}; \boldsymbol{\theta}_c^{(n-1)}; \{p_{ij}^{(n-1)}\}; s_0\right)$;

(c) backcasting: $\tilde{p}_{i,0}^{(n)} = \Pr\left(s_0 = i \mid \mathbf{Y}_T; \tilde{\boldsymbol{\lambda}}^{(n-1)}\right)$.

It can be shown that the estimates in (i) and (ii) converge to the ML estimator of the parameters of the MS-VAR. All the computations reported in this paper were carried with the MSVAR class for Ox (see Krolzig (1998), and Doornik (2001)).

5.3 Markov-Switching Models of the Euro-Zone Business Cycle

5.3.1 A Three-Regime Model of Real GDP Growth (1973-2002)

We start by applying the MSVAR approach to the disaggregated real GDP growth in the euro-zone. The quarterly, seasonally adjusted data from 1973Q3 to 2002Q1 are drawn from the OECD Main Economic Indicators data base. Since cointegration analysis gave no clear indication of the presence of a cointegrating vector, all variables are modelled in first differences.

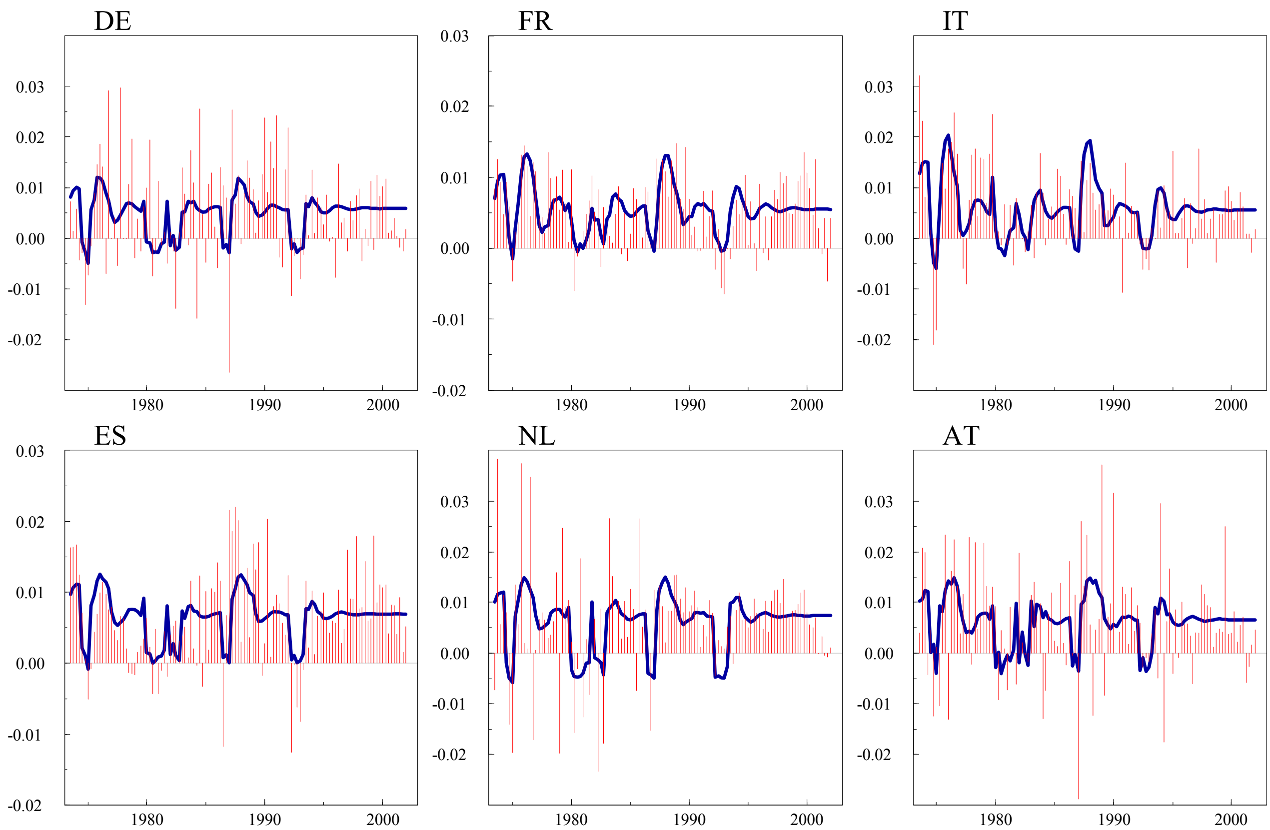

Figure 12: GDP growth data (OECD) and their common cycle

Thus, $\Delta\mathbf{y}_t$ is the vector of GDP growth rates of Germany (DE), Finland (FI), France (FR), Italy (IT), Spain (ES), the Netherlands (NL) and Austria (AT). Given the findings in the literature — particularly the observation made in (Krolzig and Toro 2000) and (Krolzig 2001a) that the catching-up process of the Mediterranean countries in the 1970s constitutes a third regime, which is fundamentally different from the traditional business cycle concept — we consider a three-regime MS-VAR model with regime-dependent intercepts and variance:

$$\Delta\mathbf{y}_t = \boldsymbol{\nu}(s_t) + \mathbf{A}_1\Delta\mathbf{y}_{t-1} + \mathbf{A}_4\Delta\mathbf{y}_{t-4} + \mathbf{u}_t, \quad \mathbf{u}_t|s_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma}(s_t)). \tag{22}$$

According to the mnemonic proposed in §2.2, the model is denoted MSIH(3)-VAR(4). The maximum lag-order is four; the autoregressive parameters at lag 2 and 3 were found to be insignificant and have been dropped from the model. Graphs of the time series reporting real GDP growth for the six euro-zone countries under consideration can be found in Figure 12. Even after the outlier correction with *PcGets* (see Hendry and Krolzig (2001)), there are still some possible outliers remaining, which could affect the ability of the MS-VAR to extract the common business cycle component from the variables.

The estimated parameters are presented in Table 2. Recessions are more pronounced in Germany, the Netherlands and Austria. Interestingly, these countries had *de facto* a monetary union for almost the entire sample period, with the Austrian and Dutch central bank shadowing the Bundesbank's monetary policy and the exchange rates moving in the narrowest of bands.

One might wonder whether this indicates an increase in business cycle symmetry in monetary unions. All countries exhibit higher growth rates in the third regime, which can be thought of as a 'high-growth regime'. For most countries it is also associated with an increased volatility of growth. This, however, is not the case for Germany where the standard errors appear to be regime-invariant.

The matrix of estimated transition probabilities is given by:

$$[\tilde{p}_{ij}] = \begin{bmatrix} 0.768 & 0.136 & 0.096 \\ 0.042 & 0.943 & 0.016 \\ 0.096 & 0.101 & 0.803 \end{bmatrix},$$

where $p_{ij} = \Pr(s_t = j|s_{t-1} = i)$.

From the estimated transition probabilities, the measures of the persistence of the regimes can be deduced, which are summarized in Table 3.

The expected number of quarters a recession prevails (duration) and the unconditional (ergodic) probability of recessions. Whereas recessions (regime 1) have an expected duration of 4 quarters, regimes 2 and 3 prevails 17 and 5 quarters, respectively. The unconditional probability of a recession is 18%.

The contribution of the common business cycle to the process of economic growth in the six euro-zone countries is depicted in Figure 12. Major differences in the mean growth rate across regimes are evident. But there are also differences with regard to the countries. For example, some countries recover faster from the 1992/93 recession than others. Particularly, the German economy appears to be falling behind its historical record of economic growth in expansions. Overall, the euro-zone cycle is most closely matched by the Dutch growth rates.

Table 2: ML estimate of the MSIH(3)-VAR(4) of real GDP growth (1973Q3 – 2002Q1)

	DE	FR	IT	ES	NL	AT
Mean growth rate (intercept)						
Recession	-0.122	0.148	0.142	0.065	-0.265	-0.098
Expansion	0.666	0.491	0.617	0.697	0.906	0.871
High growth	0.861	0.616	1.354	0.947	1.117	1.159
Short-run dynamics						
DE	0.053	-0.092	0.147	-0.094	-0.220	-0.112
FR	0.044	-0.195	-0.368	0.141	-0.354	0.805
ES	0.251	0.557	-0.019	0.243	0.399	0.639
IT	-0.276	0.031	0.124	-0.014	-0.057	-0.916
NL	-0.041	-0.074	-0.020	-0.070	-0.284	-0.114
AT	-0.076	-0.078	-0.002	-0.155	0.282	-0.353
Standard errors						
Recession	0.832	0.329	0.588	0.671	0.983	0.612
Expansion	0.851	0.434	0.592	0.539	0.630	0.820
High growth	0.765	0.503	0.817	0.604	1.395	1.386

The coefficients of the short-run dynamics report $\hat{A}_1 + \hat{A}_4$.

5.3.2 An MS-VAR of Growth in Industrial Production (1974-2002)

Next we check if we can confirm our previous findings when macroeconomic activity is measured by indices of industrial production (IIP) for eight EMU member states: Germany (DE), France (FR), Spain (ES), Belgium (BE), Austria (AT), Portugal (PT), Greece (GR) and Luxembourg (LU). The sample period again is 1974Q1 to 2002Q1. The data drawn from the OECD database are plotted in Figure 13.

The time series exhibit a more erratic behavior than the GDP growth rates, which might indicate the presence of outliers. It could also be due to moving-average errors, which have been dealt with in (Artis, Krolzig, and Toro 2004) by pre-smoothing.

In contrast to (Artis, Krolzig, and Toro 2004), we failed in finding a meaningful three-regime model of the euro-zone business cycle. Interestingly, the problems in modelling euro-zone IIP growth are also reflected in (Anas and Ferrara 2002), who analyze aggregated IIP data for the euro-zone.

Table 3: Regime properties of the MSIH(3)-VAR(4) of real GDP growth (1973Q3 – 2002Q1)

	observations	ergodic probability	expected duration
Regime 1	22.0	0.181	4.32
Regime 2	73.0	0.678	17.43
Regime 3	20.0	0.141	5.07

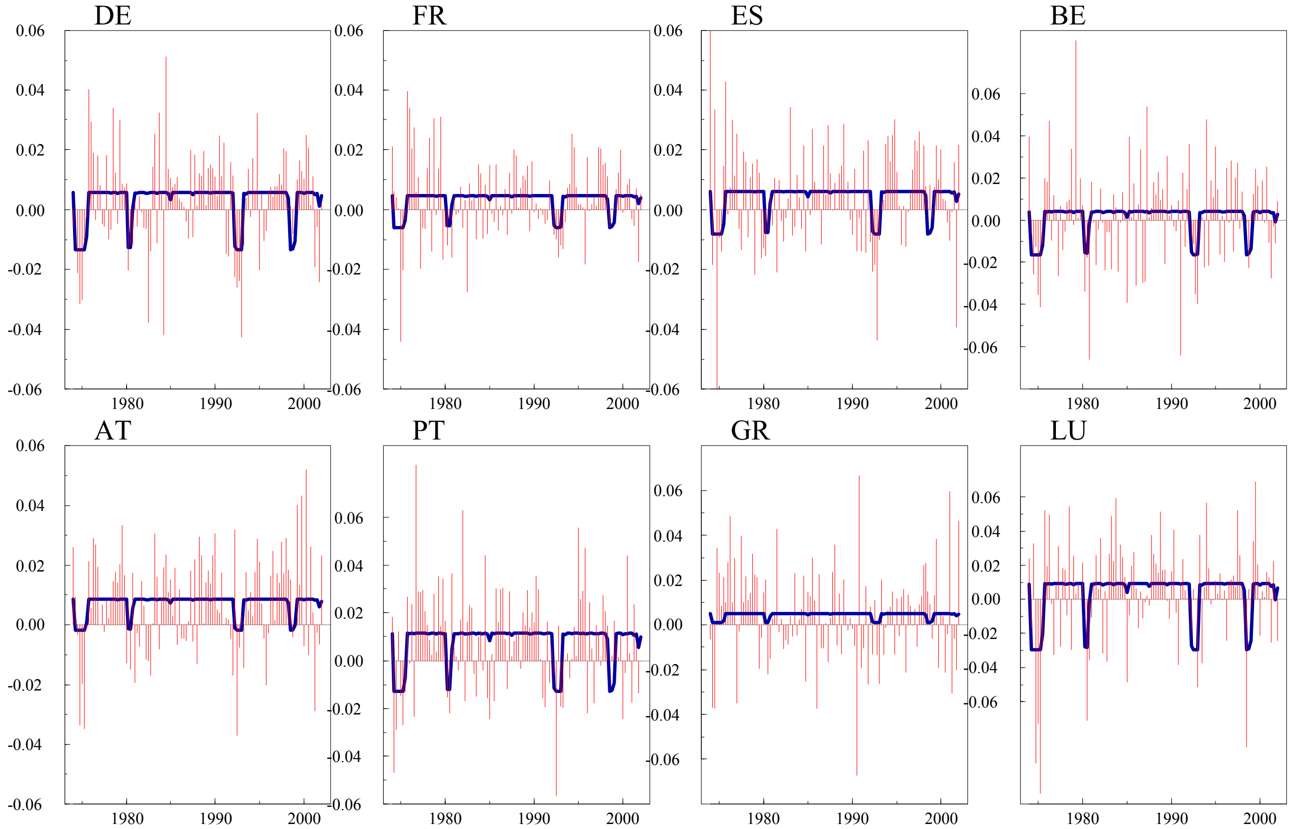

Figure 13: IIP growth data (OECD) and their common cycle

We therefore invoke the classical two-regime model. An MSM(2)-VAR(1) model with shifts in the mean growth rate of industrial production is considered:

$$\Delta \mathbf{y}_t - \boldsymbol{\mu}(s_t) = A_1 \left(\Delta \mathbf{y}_{t-1} - \boldsymbol{\mu}(s_{t-1})\right) + \mathbf{u}_t, \quad \mathbf{u}_t | s_t \sim \mathsf{NID}(\mathbf{0}, \Sigma), \tag{23}$$

where $\Delta \mathbf{y}_t$ is an (8×1) vector of IIP growth rates and $s_t \in \{1, 2\}$ is again generated by a hidden Markov chain. The regimes are distinguished by the vectors $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ defining the mean growth rates of $\Delta \mathbf{y}_t$ in '*contractions*' and '*expansions*'.

The estimated parameters are displayed in Table 4. Regime 1 is associated with contractions in industrial production in all considered countries but Greece, where output stagnates. The transition probabilities are estimated as:

$$[\tilde{p}_{ij}] = \begin{bmatrix} 0.709 & 0.291 \\ 0.047 & 0.953 \end{bmatrix},$$

resulting in the regime properties reported in Table 5. The persistence of recessions is in line with the results for MS-VAR model of GDP growth.

Table 4: ML estimate of the MSM(2)-VAR(2) in industrial production ($1974Q1 - 2002Q1$)

	DE	FR	ES	BE	AT	PT	GR	LU
				Mean growth rate				
Recession	-1.341	-0.583	-0.812	-1.645	-0.184	-1.287	0.092	-2.979
Expansion	0.576	0.464	0.618	0.414	0.849	1.140	0.492	0.891
				Short-run dynamics				
DE	-0.149	0.192	-0.101	-0.020	0.374	-0.306	0.347	0.002
FR	-0.015	-0.282	-0.254	0.361	0.178	0.189	0.432	0.271
ES	0.238	0.385	0.322	0.258	0.225	0.223	-0.198	0.137
BE	0.005	0.052	-0.041	-0.356	-0.061	-0.091	-0.189	0.116
AT	0.362	0.036	0.251	0.444	-0.105	0.289	-0.254	0.258
PT	-0.064	-0.086	-0.114	-0.071	0.029	-0.037	0.047	-0.024
GR	0.189	0.190	0.368	0.043	0.144	0.161	-0.079	-0.029
LU	-0.084	0.080	-0.081	-0.024	0.107	-0.027	0.041	-0.147
				Standard errors				
σ	1.321	0.956	1.337	1.915	1.338	1.881	1.888	2.503

The coefficients of the short-run dynamics report $\hat{A}_1 + \hat{A}_2$.

Visual inspection of the smoothed probabilities of recessions plotted in Figure 15 reveals a much shorter recession in the 1980s, no recessionary tendencies in 1986/87, and strong indication of a recession in 1998/99. Due to the pronounced difference to the regime classification produced by the previous model and the limited data quality, we have investigated the robustness of the results by the removing Luxembourg (LU) from the system. The revised model fails in detecting the recession in the early eighties. Recessions are called for the periods $1974Q1 - 1975Q2$, $1992Q2 - 1993Q2$ and finally for $2001Q2 - 2001Q4$. The lack of robustness is a reason for concern. The results derived from this model, therefore, should be considered with some caution.

5.3.3 A Two-Regime Model of GDP Growth (1980-2002)

As the third model we analyze GDP growth data provided by Eurostat over the period $1980Q3$ – $2002Q1$ plotted in Figure 14. The data are seasonally adjusted by X-12-ARIMA and outlier corrected using PcGets. We are again interested in detecting the shifts in the mean growth rate of real DGP representing turning-points of the euro-zone business cycle. The cycle is characterized by smoother transitions than in the IIP case, so we use a model with switching intercepts. Given the evidence in the literature, it can be assumed that for sample period considered here, two-regime models are an adequate description of the business cycle.

Table 5: Regime properties of the MSM(2)-VAR(2) in industrial production ($1974Q1 - 2002Q1$)

	observations	ergodic probability	expected duration
Regime 1	15.5	0.138	3.44
Regime 2	97.5	0.862	21.39

Table 6: ML estimate of the MSI(2)-VAR(1) of real GDP growth (1980Q3 – 2002Q1)

	DE	FR	IT	ES	NL	BE	FI
			Mean growth rate (intercept)				
Recession	0.162	0.245	0.020	0.220	-0.051	0.049	-0.400
Expansion	0.734	0.568	0.647	0.874	1.008	0.532	0.895
			Short-run dynamics				
DE	-0.056	-0.155	-0.011	-0.245	0.316	0.124	-0.326
FR	-0.079	-0.232	0.036	0.183	0.435	0.064	0.251
IT	0.244	0.124	-0.188	0.065	-0.106	0.112	-0.111
ES	0.220	0.206	0.100	-0.100	-0.018	0.097	0.282
NL	-0.117	-0.090	-0.053	-0.153	-0.344	-0.081	-0.079
BE	-0.031	0.241	0.137	0.300	-0.243	-0.168	0.156
FI	-0.335	-0.061	-0.063	0.018	-0.135	-0.121	-0.068
			Standard errors				
σ	0.750	0.522	0.537	0.652	0.852	0.777	0.897

A first-order model with shifts in the intercept is used:

$$\Delta \mathbf{y}_t = \boldsymbol{\nu}(s_t) + \mathbf{A}_1 \Delta \mathbf{y}_{t-1} + \mathbf{u}_t, \quad \mathbf{u}_t | s_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma}). \tag{24}$$

The ML estimates of this MSI(2)-VAR(1) model are reported in Table 6.The estimated transition probabilities are given by:

$$[\tilde{p}_{ij}] = \begin{bmatrix} 0.898 & 0.102 \\ 0.053 & 0.947 \end{bmatrix},$$

implying the regime properties reported in Table 7.

Only in the case of Finland and the Netherlands, real GDP contracts in the first regime. Figure 14 reveals that the very severe recession in Finland during the early nineties is indeed the mostly remarkable feature of the data set. In contrast, the Markov component of GDP growth is positive for all other euro-zone countries over the entire sample period. It is therefore more appropriate to think of this model as a representation of the euro-zone *growth* cycle. This finding is confirmed by the surprisingly high estimate of the expected duration of the '*low-growth*' regime. The average duration of a regime 1 is with 10 quarters twice as long as the 'euro-zone recessions' found by the first two models. However, the estimated duration is very similar to the result in Anas and Ferrara (2002), when analyzing aggregated euro-zone GDP data. Also, the probabilities of the '*low-growth*' regime follow closely their results (see Figure 15). Apart from the outlier in 1986Q1, there are three distinguished recessions within the sample period: 1981Q1 – 1983Q1, 1990Q2 – 1993Q1, and 2001Q2 – 2002Q2.

Table 7: Regime properties of the MSI(2)-VAR(1) of real GDP growth (1980Q3 – 2002Q1)

	observations	ergodic probability	expected duration
Regime 1	24.8	0.3405	9.78
Regime 2	62.2	0.6595	18.93

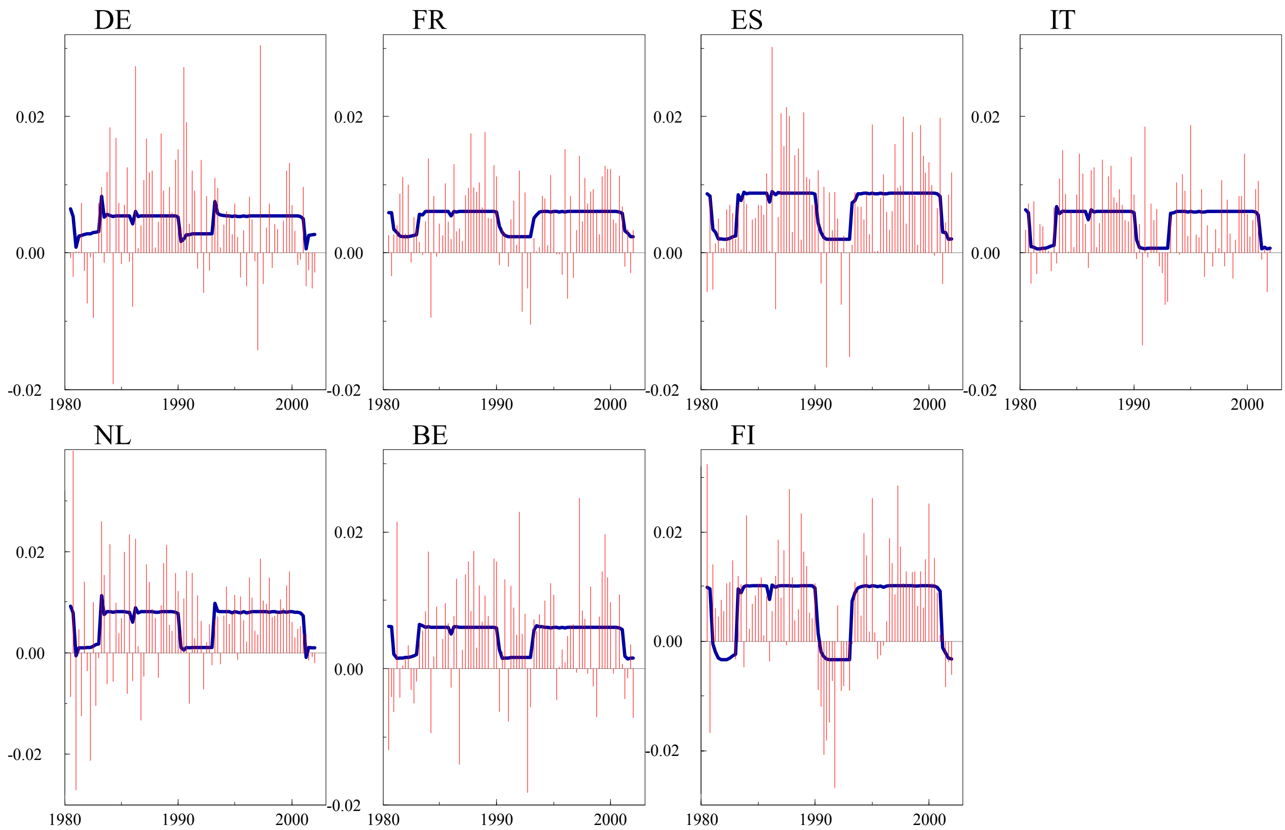

Figure 14: GDP growth data (Eurostat) and their common cycle

In § 5.5, we will compare the regime classifications derived here to the business cycle chronologies proposed in the literature. Next, we investigate the degree of business cycle synchronization in the three estimated models of the euro-zone business cycle.

5.4 Business Cycle Synchronization

The synchronization of the business cycles in the EMU member states is an important indicator for the optimality of a single monetary policy.

A lack of business cycle synchronization within the euro area could complicate the operation of monetary policy in the union. For non-member states, it constitutes a negative indicator for joining EMU. Therefore, business cycle synchronism deserves careful screening.

Here we analyze whether the cyclical shifts in the mean growth rate of the euro-zone countries are traceable to the exposed common euro-zone cycle.

The hypothesis of a common euro-zone cycle is investigated by testing for the significance of the regime shifts in the mean growth rate of the economies in the euro-zone. Under the null hypothesis, the mean growth rate of country k in recessions (regime 1) is identical to its growth rate in expansions (regime 2): $\mu_{k1} = \mu_{k2}$ for $k = 1, \ldots, K$.

Since the regime-dependent means in the remaining equations of the system are unrestricted, the model is identified under the null hypothesis. Hence, the tests are nuisance parameter free and classical likelihood theory can be invoked.

The asymptotic null distribution of the Wald test is $\chi^2(r)$, where $r = 1$ is the number of linearly independent restrictions.

Since testing for the number of regimes is a difficult enterprise in Markov-switching models, it is important to note that we are here only interested in showing that individual euro-zone countries share the inferred cycle as episodes of economic expansion and contraction (or stagnation). In contrast, formal tests of the Markov-switching model against linear alternatives are confronted with unidentified nuisance parameters under the null, hampering the applicability of conventional testing procedures. Suppose a two-regime model with Markov-switching means is subjected to the test hypothesis $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$.

Then, the transition probabilities p_{12} and p_{21} are not identified. Standardized likelihood ratio (LR) tests designed to deliver (asymptotically) valid inference have been proposed by Garcia (1998).

Hansen's approach delivers a bound on the asymptotic distribution of the standardized LR test, but the test is conservative, tending to be under-sized in practice and of low power, and computationally demanding. In comparison, the tests presented here can be easily implemented (see Krolzig (1997b), ch.7, for details) and were found useful in empirical research.

Table 8 reports the results of the Wald specification tests for the regime invariance of the mean growth rates for the three models discussed in this paper.

Table 8: Wald tests of the regime-invariance hypothesis

Model	MSIH(3)-VAR(4)						MSM(2)-VAR(2)		MSI(2)-VAR(1)	
Data	GDP growth						IIP growth		GDP growth	
Source	OECD						OECD		Eurostat	
Sample	1973Q3-2002Q1						1974Q1-2002Q1		1980Q3-2002Q1	
Test	A	$\chi^2(1)$	B	$\chi^2(1)$	C	$\chi^2(2)$	A	$\chi^2(1)$	B	$\chi^2(1)$
DE	13.62	0.02	0.82	36.46	16.27	0.03	11.79	0.06	5.70	1.69
FR	13.16	0.03	0.93	33.32	16.07	0.03	7.51	0.61	4.23	3.95
IT	14.04	0.02	2.36	12.39	17.71	0.01			14.77	0.01
ES	9.32	0.23	13.14	0.03	25.94	0.00	9.60	0.19	11.03	0.09
NL	24.43	0.00	0.39	53.00	25.05	0.00			15.75	0.01
BE							2.30	12.91	22.04	0.00
AT	28.07	0.00	0.69	40.34	30.39	0.00	3.68	5.50		
FI									4.02	4.48
PT							4.80	2.83		
GR							0.38	53.34		
LU							8.93	0.28		
System	180.21	0.00	57.32	0.00	659.50	0.00				

The reported system tests are free of nuisance parameters since the regimes are identified under the null by switching covariance matrices $\Sigma_i \neq \Sigma_j$ for $i \neq j$. The degree of freedoms are 6 and 12, respectively. m.s.l. of the χ^2 statistics have been multiplied by 100. A states for the test for $\mu_1 = \mu_2$, B states for the test for $\mu_2 = \mu_3$ and C states for the test for $\mu_1 = \mu_2 = \mu_3$.

In particular:

i. MSIH(3)-VAR(4) of GDP growth (OECD data): the test hypothesis $\nu_{k1} = \nu_{k2}$ is strongly rejected for all countries and the system as a whole. However, the support for a third regime is not very strong. Only in the case of Spain (ES), the differences in the regime-dependent intercept turn out to be statistically significant.[6] Due to the regime dependence of the variance matrices, also system tests of the form $\boldsymbol{\nu}_1 = \boldsymbol{\nu}_2 = \boldsymbol{\nu}_3$ are free of nuisance parameters. The test results provide strong support for the presence of regime shifts in the mean growth rate of the analyzed economies.

ii. MSM(2)-VAR(2) of IIP growth (OECD data): the test hypothesis of regime-invariant means, $\mu_{k1} = \mu_{k2}$, is strongly rejected for Germany (DE), France (FR), Spain (ES) and Luxembourg (LU). It can also be rejected for Portugal (PT), with Austria (AT) being a marginal case. Belgium (BE) and Greece (GR), however, are not significantly affected by the euro-zone *growth* cycle.

iii. MSI(2)-VAR(1) of GDP growth (Eurostat data): the regime-invariance hypothesis $\mu_{k1} = \mu_{k2}$ can be rejected for all countries, though the common cycle is less pronounced in Germany (DE), France (FR) and Austria (AT).

Overall, the notion of a common euro-zone business cycle is justified.There is overwhelming evidence for the relevance of the identified euro-zone *growth* and *business* cycles. With the exception of Greece, the common cycle contributes to the mean growth rate in all countries of the euro-zone.

5.5 Turning Point Chronologies of the Euro-Zone Business Cycle

As discussed in § 3, the construction of a turning point chronology with Markov-switching models is based on the time paths of the smoothed regime probabilities, which are a by-product of the EM algorithm. For the dating of the euro-zone business cycle, we start by reconsidering the smoothed and filtered regime probabilities derived with the three models considered above.

The regime probabilities are plotted in Figure 15; the filtered ones are shown with bars, and the smooth ones are represented by a bold line. It is worth recalling that the filtered regime probability can be understood as an optimal inference on the state variable (whether the system is in a boom or recession) at time t using only the information available up to this point in time, i.e. $\Pr(s_t = m \mid \mathbf{Y}_t)$, where m stands for a given regime; the smoothed regime probability is the optimal inference on the regime at time t using the full sample information Y_T: $\Pr(s_t = m \mid \mathbf{Y}_T)$. The difference between the two measures indicates the amount of revisions required when new observations become available.

The classification of the regimes and the dating of the business cycle amounts to assigning $\mathbf{y}_t$ to the regime m with the highest smoothed probability. The resulting chronology of the euro-zone business cycle is reported in Table 9 for the three measures of macroeconomic activity analyzed

[6] Note, however, that regime 3 is also characterized by a sharply increased volatility for Italy (IT), the Netherlands (NL) and Austria (AT).

in this paper. The peak P date denotes the period t just before the beginning of a recession, i.e. Pr(‘*recession*’ at time $t \mid \mathbf{Y}_T) < 0.5$ and Pr(‘*recession*’ at time $t+1 \mid \mathbf{Y}_T) > 0.5$; the trough T is the last period of the recession.

Table 9: Datings of the euro-zone business cycle

Model	MSIH(3)-VAR(4)			MSM(2)-VAR(2)			MSI(2)-VAR(1)		
Data	GDP growth			IIP growth			GDP growth		
Source	OECD			OECD			Eurostat		
Sample	1973Q3-2002Q1			1974Q1-2002Q1			1980Q3-2002Q1		
Cycle	Peak	Trough	h	Peak	Trough	h	Peak	Trough	h
A	1974Q2	1975Q1	0.75	*1974Q1*	1975Q3	1.50			
B									
C	1979Q4	1981Q3	1.75	1980Q1	1980Q3	0.50	1980Q4	-	
D	1981Q4	1982Q4	1.00				-	1983Q1	2.25
E	1986Q2	1987Q1	0.75						
F	1992Q1	1993Q3	1.50	1992Q1	1993Q1	1.00	1990Q1	1993Q1	3.00
G				1998Q2	1999Q1	0.75			
H							2001Q1	[2002Q2]	1.25

In italics period before the beginning or at the end of the sample; h Duration of recessions in years.

Table 10: Datings of the euro-zone business cycle in the literature

	K1			K2†			A1			A2		
Cycle	P	T	h	P	T	h	P	T	h	P	T	h
A										74M8	75M8	1.00
B										77M2	77M7	0.41
C	*80Q1*	81Q1	1.00	*80Q1*	81Q1	1.00	*80Q1*	82Q4	3.00	80M4	80M10	0.50
D										82M5	82M8	0.25
E												
F	92Q1	93Q1	1.00	92Q2	93Q3	1.25	92Q1	93Q2	1.00	92M4	93M2	0.83
G												
H							01Q1	02Q2	1.00	01M10	01M12	0.17

	AKT			KT*			PS		
Cycle	P	T	h	P	T	h	P	T	h
A	74M7	75M7	1.00	74Q1	75Q2	1.25			
B									
C	79M10	-		80Q1	-		79Q4	-	
D	-	82M8	2.83	-	82Q4	2.75	-	83Q1	4.25
E							85Q4	87Q1	1.25
F	90M9	92M9	2.00	92Q2	93Q2	1.00	90Q1	92Q3	2.50
G							95Q2	96Q1	0.75
H									

In italics period before the beginning or at the end of the sample; h Duration of recessions in years; * Dummy variables being 1 in 1984Q2 and 1987Q1 and -1 in the subsequent period included; Dummy variables being 1 in 1987Q1 and 1987Q1 and -1 in the subsequent period included.

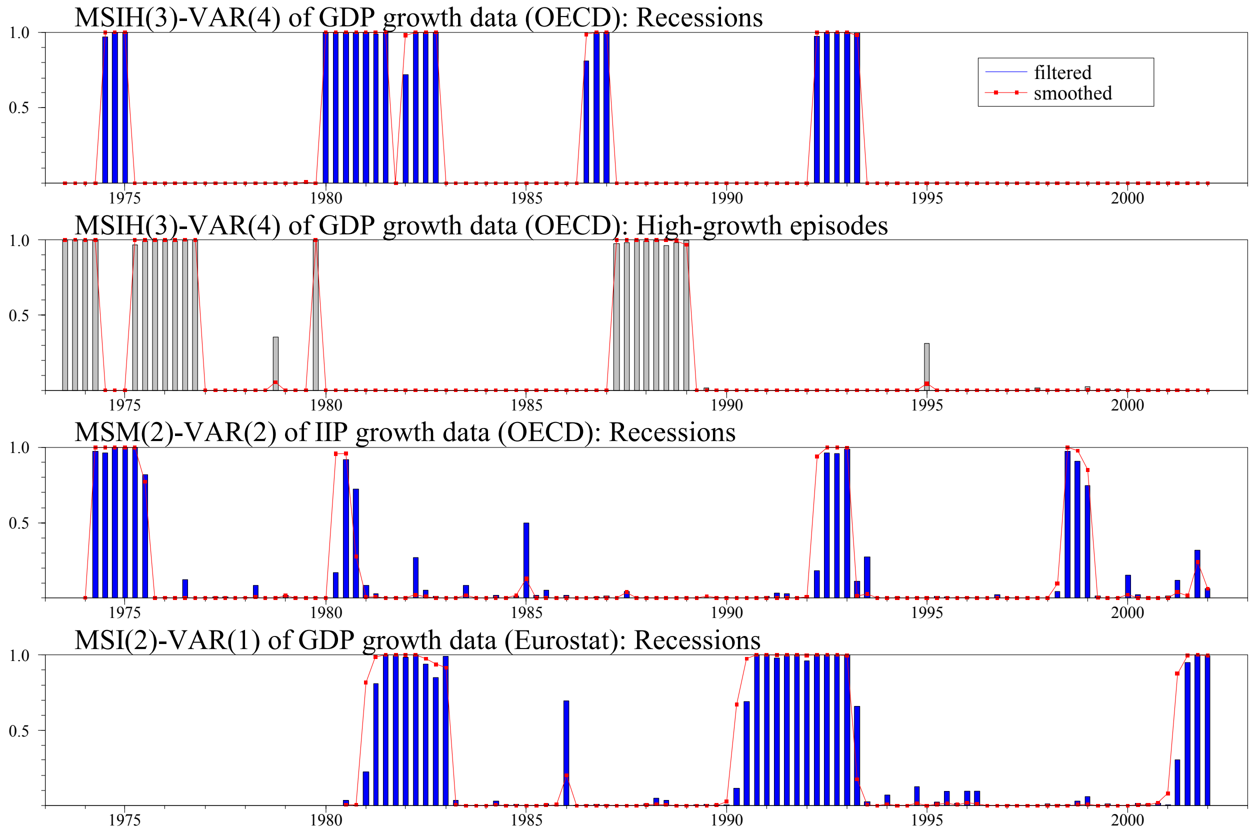

Figure 15: Regime probabilities

The derived business cycle chronologies should be compared with the datings in Table 10, which reports the datings of the European business cycle derived from Markov-switching models proposed in the literature (see the overview in Table 1 of § 5.1).

Given the lack of an NBER-type reference cycle for the euro zone, it is worth to elaborate the robust features of the euro-cycle datings in the literature and to compare them to the turning points of the euro-zone business cycle identified in this paper. The most striking result is that the GDP data based business cycle classifications of Krolzig (2001b) and Anas and Ferrara (2002) coincide when the different sample length is taken into consideration.

In contrast, the MS-AR mode considered by Anas and Ferrara (2002) for euro-zone IIPgenerates recessions which are too short and too frequent.

The main difference with regard to the results of Artis, Krolzig, and Toro (2004), Krolzig and Toro (2000) and Krolzig (2001a)} is the relative short expected duration of recessions. As these papers find some indication for incomplete business cycle synchronization, the prolonged duration of recessions may reflect the need to subsume the UK and continental cycle.

An outlier is marked by the study of Peersman and Smets (2001) as the common cycle identified in their study has more the characteristics of a *growth* cycle. As can be seen in Table 10, their (growth) ‘*recessions*’ have much a longer duration and are found much more frequently.

Altogether, we can conclude that a robust pattern in the chronologies of the euro-zone business cycle produced by Markov-switching models has emerged.

The turning point chronology of the euro-zone business cycle is dominated by three recessions:

i. the recession following the 1973/74 oil price shock (from 1974Q2/Q3 to 1975Q1/Q3);

ii. possibly a double dip recession in the early eighties (1980Q1/1981Q1 to 1982Q4/1983Q1), coinciding with Volcker disinflation in the US;

iii. a recession in the early nineties (1990Q2/1992Q2 to 1993Q1/Q3) associated with greater uncertainty regarding its dating, presumably caused by monetary tightening and fiscal expansion in the aftermath of the German reunification;

iv. there is some indication of a euro-zone recession commencing in 2001Q2

Some differences in the precise dating of the turning points remain. Among other causes, they are likely due to the choice of the measure of macroeconomic activity and the set of countries, for which the data have been available. In line with the literature[7], we have identified various characteristics of the data that matter for the construction of the business cycle chronology:

i. Provision of data for the largest possible set of countries;

ii. Availability of long historic time series;

iii. Data frequency;

iv. Pre-filtering:

- seasonal adjustment;
- outlier correction;
- smoothing (for monthly data sets).

The importance of the quality of the data for the dating of the euro-zone business cycle underlines the pivotal role of Eurostat in the creation and distribution of high-quality statistical data for the EMU member states. The case for further research is clearly granted.

[7]For example, it is well known that seasonal adjustment procedures can severely affect the determination of business cycle turning points (see, *inter alia*, Astolfi, Ladiray, and Mazzi (2001), and Lommatzsch and Stephan (2001)).

6 Conclusions

In this paper, we have addressed the issue whether Markov-switching vector autoregressive models can be used for the extraction and dating of the euro-zone business cycle. We have shown that this generalization of the approach innovated by Hamilton in his analysis of the US business cycle allows the identification of a common business cycle in the euro zone. The regime identification scheme distinguished between states of recession and expansion in the euro area, where different expansionary regimes may be considered to embed structural change.

The main features of the proposed MS-VAR modelling procedure are:

i. *Approach*

- Congruent statistical model of the time series of interest;
- Meaningful definition of a common (euro-zone) business cycle.

ii. *Business chronologies*

- Dating of turning point based on the smoothed probability of being in a recession;
- Multi-regime models feasible.

iii. *Advantages of the regime-switching approach*

- Quantification of the uncertainty associated with the regime classification;
- Detection of new turning points in real time;
- Systematic revision when new data arrive.

iv. *Advantages of modelling systems*

- Euro-zone business cycle as a common feature of the macroeconomic dynamics in the EMU member states;
- Missing invariance of regime classification under aggregation.

The empirical analysis consisted in fitting an MS-VAR model to a system of real GDP or IIP growth rates of EMU member states. The models have been found statistically congruent and economically meaningful, though room for improvements and refinements remains. The results shed light on the issue of which countries drive the euro-zone cycle. We have been able to demonstrate the relevance of the common cycle for all EMU member states except Greece.

So the evidence for the presence of a common, Markovian euro-zone business cycle shared by the EMU member states is very robust. The estimated Markov-switching VAR models were then used to construct turning point chronologies of the euro-zone business cycle, differing regarding the measurement of macroeconomic activity, the set of countries available, and the time period considered. While the pattern of recessions and expansions is robust, differences in the precise dating of events have been acknowledged. In this context, we have also emphasized the importance of data quality. Altogether, we believe that the findings of this study justify a recommendation of the Markov-switching vector autoregressive model for the construction of turning point chronologies of the euro-zone business cycle.

References

Anas, J., and L. Ferrara (2002): "A comparative assessment of parametric and non parametric turning points detection methods: the case of the Euro-zone economy," Discussion Paper, Centre d'Observation Economique, Paris.

Artis, M., H.-M. Krolzig, and J. Toro (2004): "The European Business Cycle," *Oxford Economic Papers*, 56.

Astolfi, R., D. Ladiray, and G. L. Mazzi (2001): "Business Cycle Extraction of Euro-Zone GDP: Direct versus Indirect Approach," *Vierteljahreshefte zur Wirtschaftsforschung*, 70, 377–398.

Baum, L. E., and T. Petrie (1966): "Statistical inference for probabilistic functions of finite state Markov Chains," *Annals of Mathematical Statistics*, 37, 1554–1563.

Burns, A. F., and W. C. Mitchell (1946): *Measuring Business Cycles*. NBER, New York.

Diebold, F. X., J. H. Lee, and G. C. Weinbach (1994): "Regime Switching with Time-Varying Transition Probabilities," in *Non-stationary Time-series Analyses and Cointegration*, ed. by C. Hargreaves, pp. 283–302. Oxford University Press, Oxford.

Doornik, J. A. (2001): *Object-Oriented Matrix Programming using Ox*. Timberlake Consultants Press, London, 4th edn.

Filardo, A. J., and S. F. Gordon (1994): "International Co-Movements of Business Cycles," Research Working Paper 94-11, Federal Reserve Bank of Kansas.

Garcia, R. (1998): "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," *International Economic Review*, 39.

Hamilton, J. D. (1989): "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, 57, 357–384.

——— (1990): "Analysis of time series subject to changes in regime," *Journal of Econometrics*, 45, 39–70.

Hamilton, J. D., and G. Lin (1996): "Stock Market Volatility and the Business Cycle," *Journal of Applied Econometrics*, 11, 573–593.

Hendry, D. F., and H.-M. Krolzig (2001): *Automatic Econometric Model Selection with PcGets*. Timberlake Consultants Press, London.

Kim, C.-J. (1994): "Dynamic linear models with Markov-switching," *Journal of Econometrics*, 60, 1–22.

Krolzig, H.-M. (1996): "Statistical Analysis of Cointegrated VAR Processes with Markovian Regime Shifts," SFB 373 Discussion Paper 25/1996, Humboldt Universität zu Berlin.

——— (1997a): "International Business Cycles: Regime Shifts in the Stochastic Process of Economic Growth," Applied Economics Discussion Paper 194, Department of Economics, University of Oxford.

——— (1997b): *Markov Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis*. Springer, Berlin.

——— (1998): "Econometric Modelling of Markov-Switching Vector Autoregressions using MSVAR for Ox," Discussion Paper, Department of Economics, University of Oxford: http://www.econ.ox.ac.uk/research/hendry/krolzig.

——— (2001a): "Business Cycle Measurement in the Presence of Structural Change: International Evidence," *International Journal of Forecasting*, 17, 349–368.

——— (2001b): "Markov-Switching Procedures for Dating the Euro-zone Business Cycle," *Vierteljahreshefte zur Wirtschaftsforschung*, 70, 339–351.

——— (2003a): "Business cycle analysis and aggregation. Results for Markov-switching VAR processes," Discussion Paper, Department of Economics, University of Oxford.

——— (2003b): "Predicting Markov-Switching Vector Autoregressive Processes," *Journal of Forecasting*, 56, forthcoming.

Krolzig, H.-M., and J. Toro (2000): "Classical and Modern Business Cycle Measurement: The European Case," Discussion Paper in Economics 60, University of Oxford.

Lindgren, G. (1978): "Markov Regime Models for Mixed Distributions and Switching Regressions," *Scandinavian Journal of Statistics*, 5, 81–91.

Lommatzsch, K., and S. Stephan (2001): "Seasonal Adjustment Methods and the Determination of Turning Points of the EMU Business Cycle," *Vierteljahreshefte zur Wirtschaftsforschung*, 70, 399–415.

Peersman, G., and F. Smets (2001): "Are the effects of monetary policy in the Euro area greater than in booms," ECB Working Paper 52, European Central Bank, Frankfurt.

Phillips, K. (1991): "A Two-Country Model of Stochastic Output with changes in regime," *Journal of International Economics*, 31, 121–142.

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Cyclical Convergence and Forecasting"

CONVERGENCE AND CYCLES IN THE EURO-ZONE

By Vasco M. Carvalho and Andrew C. Harvey[1]

University of Cambridge

Multivariate unobserved components (structural) time series models are fitted to annual post-war observations on real income per capita in countries in the Euro-zone. The aim is to establish stylised facts about convergence as it relates both to long-run income levels and to cycles. The analysis is based on a new model in which convergence components are combined with a common trend and similar cycles. These convergence components are formulated as a second-order error correction mechanism which ensures that the extracted components change smoothly thereby giving a clearer decomposition into long-run movements and cycles.

Keywords: Balanced growth; Error correction mechanism; Kalman filter; Signal extraction; Stochastic trend; Unobserved components.

JEL classification: C32; O40.

[1]The work was supported by the Economic and Social Research Council (ESRC) as part of a project on Dynamic Common Factor Models for Regional Time Series, grant number L138 25 1008. An earlier version was presented at a conference on Modern Tools for Business Cycle Analysis, Eurostat, Luxembourg, November 2002.

1 Introduction

Minimising income inequality across its members has long been a declared objective of the European Union. In particular, promoting convergence of income in levels (or at least the conditions for that to be obtained) has been the justification for the establishment first, of Structural Funds (the largest of which the European Regional Development Fund established in 1975) and, more recently (1993), of the Cohesion Fund. The latter is specifically directed at poorer countries (Greece, Ireland, Portugal and Spain) and was set up with the purpose of making compatible the EMU budgetary discipline requirements and the continued infrastructure investment requirements in these countries, which the EU deems as necessary for convergence to take place. As Boldrin and Canova (2001) point out, what this implies about the view underlying EU policies is that deepening economic and monetary integration, by itself, leads to divergence of income levels across Europe; see also Martin (2000). However, this view stands at odds with neo-classical growth theories. As economic and monetary integration deepens and free movement of goods, people and capital becomes a reality across Europe, the preconditions for such theories are more likely to be met. As such, the theoretical prediction would be one of convergence not divergence.

Which view is correct is not only of theoretical interest but also has important policy implications. If the EU view is correct, the creation of EMU means that incentives for structural adjustment are needed for catching up to take place.[2] Given the loss of national monetary policies and the strict budgetary requirements limiting national public spending, European-wide redistribution would be the only option. Indeed, a substantial part of EU's resources is already directed at sustaining cohesion: Boldrin and Canova (2000) find that, over the 1986-1999 period, it is close to 8% of the Community's GDP.

Given such vast policy implications, the need to verify the validity of the underlying premise is obvious. Stylised facts on growth and convergence are needed to answer questions such as: are Euro-zone economies converging to a single steady state distribution or are they clustering around different states? Were they diverging before the establishment of Structural and Cohesion funds? Is there any visible effect of the latter on the growth dynamics of poorer countries?

However the fact remains that the existing literature on European convergence - mirroring the situation throughout the entire growth empirics literature- has reached no agreement on what the European record really is. This stems mainly from the use of different methods; see Durlauf and Quah (1999) for a review and criticism of the empirical literature. Thus while the early cross-sectional approach of Barro and Sala-i-Martin (1992) concluded in favour of absolute but slow convergence in Europe, panel models with fixed effects have suggested very fast convergence towards different steady states. Finally, following the time-series/cointegration approach proposed in Bernard and Durlauf (1995, 1996), Tsionas (2000) applies a battery of unit-root/stationarity tests to conclude that both divergence and convergence are possible: 'the results are mixed and depend critically on the type of test employed'.

[2]The Delors Report of 1989 makes this clear, stating that with deeper integration 'transport costs and economies of scale would tend to favor a shift in economic activity away from less developed regions (...) to the highly developed areas (...). The economic and monetary union would have to encourage and guide structural adjustment which would help poorer regions catch up with the wealthier ones'.

As we have argued elsewhere (see Harvey and Carvalho (2002) and Carvalho and Harvey (2002)) we believe that rather than adding yet another set of regressors or testing in a vacuum, presenting the stylised facts on growth and convergence is of more value both for the development of meaningful theories and for policy decisions. This response is in line with the growing dissatisfaction on the current state of growth empirics and Durlauf's (2001, p 68) call for econometrics to 'clarify how empirical workers should elucidate data patterns and draw inferences concerning growth'. The aim of the present paper is to establish stylised facts about convergence in Euro-zone countries both with respect to long-run income levels and to cycles. Distinguishing trends from cyclical movements is essential to an effective study of convergence. The analysis is based on a new multivariate unobserved components model in which convergence components are combined with a common trend and similar cycles. These convergence components are formulated as a second-order error correction mechanism which ensures that the extracted components change smoothly thereby giving a clearer decomposition into long-run movements and cycles. Furthermore the second-order mechanism is able to capture temporary divergence, something that is a feature of the Euro-zone data. Because the cross-section is relatively small, we are able to properly account for the cross-correlations across regions. The statistical treatment of the model is based on the state space form. Parameters are estimated by maximum likelihood and components are extracted by the Kalman filter and associated smoother.

The plan of the paper is as follows. In Section 2 we review the main ideas of structural time series models and show how a multivariate model handles balanced growth. Section 3 sets out the convergence model. Following the preliminary analysis of trends in per capita income in the Euro-zone countries in Section 4, this model is fitted to two groups in Section 5. The growth paths of the two groups are then compared. Section 6 examines the evidence for convergence in cycles. Section 7 assesses the extent to which unit root and co-integration tests can provide meaningful evidence on convergence, while Section 8 concludes.

2 Structural time series models and balanced growth

2.1 Univariate models

The *local linear trend* model for a set of observations, $y_t, t = 1, .., T$, consists of a stochastic trend and an irregular component, that is

$$y_t = \mu_t + \varepsilon_t, \qquad t = 1, ..., T, \tag{1}$$

where the trend, μ_t, receives shocks to both its level and slope so

$$\begin{array}{rcll} \mu_t & = & \mu_{t-1} + \beta_{t-1} + \eta_t, & \eta_t \sim NID(0, \sigma_\eta^2), \\ \beta_t & = & \beta_{t-1} + \zeta_t, & \zeta_t \sim NID(0, \sigma_\zeta^2), \end{array} \tag{2}$$

where the irregular, level and slope disturbances, ε_t,η_t and ζ_t, respectively, are mutually independent and the notation $NID\left(0, \sigma^2\right)$ denotes normally and independently distributed with mean zero and variance σ^2. If both variances σ_η^2 and σ_ζ^2 are zero, the trend is deterministic. When only σ_ζ^2 is zero, the slope is fixed and the trend reduces to a random walk with drift. Allowing σ_ζ^2 to be positive, but setting σ_η^2 to zero gives an integrated random walk trend, which when estimated tends to be relatively smooth. The model is often referred to as the '*smooth trend*' model.

The statistical treatment of unobserved component models is based on the state space form (SSF). Once a model has been put in SSF, the Kalman filter yields estimators of the components based on current and past observations. Signal extraction refers to estimation of components based on all the information in the sample. It is based on smoothing recursions which run backwards from the last observation[3]. Predictions are made by extending the Kalman filter forward. Root mean square errors (RMSEs) can be computed for all estimators and prediction or confidence intervals constructed.

The unknown variance parameters are estimated by constructing a likelihood function from the one-step ahead prediction errors, or innovations, produced by the Kalman filter. The likelihood function is maximized by an iterative procedure. The calculations can be done with the STAMP package of Koopman et al (2000). Once estimated, the fit of the model can be checked using standard time series diagnostics such as tests for residual serial correlation.

Distinguishing a long-term trend and from short-term movements is important. Short-term movements may be captured by adding a serially correlated stationary component, ψ_t, to the model. Thus

$$y_t = \mu_t + \psi_t + \varepsilon_t, \quad t = 1, ..., T \tag{3}$$

An autoregressive process is often used for ψ_t. Another possibility is the stochastic cycle

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad t = 1, ..., T, \tag{4}$$

where λ_c is frequency in radians and κ_t and κ_t^* are two mutually independent white noise disturbances with zero means and common variance σ_κ^2. The period corresponding to λ_c is $2\pi/\lambda_c$. For $0 \leq \rho < 1$, the process ψ_t is stationary with zero mean and variance $\sigma_\psi^2 = \sigma_\kappa^2/(1-\rho^2)$; see Harvey (1989, p60). The reduced form is an $ARMA(2,1)$ process in which the autoregressive part has complex roots. The complex root restriction together with the smooth trend restriction often allows a clearer separation into trend and cycle.

2.2 Multivariate models

Suppose we have N time series. Define the vector $\mathbf{y}_t = (y_{1t}, .., y_{Nt})'$ and similarly for $\boldsymbol{\mu}_t, \boldsymbol{\psi}_t$ and $\boldsymbol{\varepsilon}_t$. Then a multivariate UC model may be set up as

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\psi}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \quad t = 1, ..., T, \tag{5}$$

where $\boldsymbol{\Sigma}_\varepsilon$ is an $N \times N$ positive semi-definite matrix. The trend is

$$\begin{aligned} \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\eta) \\ \boldsymbol{\beta}_t &= \boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\zeta), \end{aligned} \tag{6}$$

[3]In a smooth trend model with a signal-noise ratio, $\sigma_\zeta^2/\sigma_\varepsilon^2$, of 1/1600, the extracted trend corresponds to the trend obtained by the Hodrick-Prescott (HP) filter for quarterly data.

With $\mathbf{\Sigma}_\eta = 0$, we get the smooth trend model. With $\mathbf{\Sigma}_\zeta = 0$, we get the random walk plus drift.

The *similar cycle* model, introduced by Harvey and Koopman (1997) is

$$\begin{bmatrix} \boldsymbol{\psi}_t \\ \boldsymbol{\psi}_t^* \end{bmatrix} = \left[\rho \begin{pmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{pmatrix} \otimes \mathbf{I}_N \right] \begin{bmatrix} \boldsymbol{\psi}_{t-1} \\ \boldsymbol{\psi}_{t-1}^* \end{bmatrix} + \begin{bmatrix} \boldsymbol{\kappa}_t \\ \boldsymbol{\kappa}_t^* \end{bmatrix}, \quad t = 1, ..., T, \tag{7}$$

where $\boldsymbol{\psi}_t$ and $\boldsymbol{\psi}_t^*$ are $N \times 1$ vectors and $\boldsymbol{\kappa}_t$ and $\boldsymbol{\kappa}_t^*$ are $N \times 1$ vectors of the disturbances such that

$$E(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t') = E(\boldsymbol{\kappa}_t^* \boldsymbol{\kappa}_t^{*'}) = \mathbf{\Sigma}_\kappa, \quad E(\boldsymbol{\kappa}_t \boldsymbol{\kappa}_t^{*'}) = \mathbf{0}, \tag{8}$$

where $\mathbf{\Sigma}_\kappa$ is an $N \times N$ covariance matrix. The model allows the disturbances to be correlated across the series. Because the damping factor and the frequency, ρ and λ_c, are the same in all series, the cycles in the different series have similar properties; in particular their movements are centred around the same period. This seems eminently reasonable if the cyclical movements all arise from a similar source such as an underlying business cycle. Furthermore, the restriction means that it is often easier to separate out trend and cycle movements when several series are jointly estimated.

The common cycle model of Vahid and Engle (1993) assumes perfect correlation between cycles. This is a very strong restriction. The recent paper by Carlino and Sill (2001) uses the methodology of Vahid and Engle (1993) to decompose series on US regions into common trends and common cycles. In Carvalho and Harvey (2002) we explain why we do not find the resulting cycles particularly plausible.

2.3 Stability and balanced growth

The *balanced growth* UC model is a special case of (5):

$$\mathbf{y}_t = \mathbf{i}\mu_t + \boldsymbol{\alpha} + \boldsymbol{\psi}_t + \boldsymbol{\varepsilon}_t, \;, \qquad t = 1, ..., T, \tag{9}$$

where μ_t is a univariate local linear trend, $\mathbf{i}$ is a vector of ones, and $\boldsymbol{\alpha}$ is an $N \times 1$ vector of constants. If μ_t is initialised with a diffuse prior, then $\boldsymbol{\alpha}$ must be subject to a constraint so it contains only $N - 1$ free parameters, for example there may be one zero entry. Note that although the levels may be different, the slopes are the same, irrespective of whether they are fixed or stochastic.

A balanced growth model implies that the series have a stable relationship over time. This means that there is a full rank $(N - 1) \times N$ matrix, $\mathbf{D}$, with no null columns and the property that $\mathbf{Di} = \mathbf{0}$, thereby rendering $\mathbf{Dy}_t$ jointly stationary. In other words it removes the common trend, μ_t. The rows of $\mathbf{D}$ may be termed *balanced growth co-integrating vectors.* Typically each row will contain a one, a minus one and zeroes elsewhere. For example, one country may be used as a benchmark.

3 Convergence models

The multivariate convergence model proposed by Carvalho and Harvey (2002) is

$$\mathbf{y}_t = \boldsymbol{\alpha}+\beta\mathbf{i}t + \boldsymbol{\mu}_t + \boldsymbol{\psi}_t+\boldsymbol{\varepsilon}_t, \quad t = 1, ..., T \tag{10}$$

with $\boldsymbol{\alpha}'\mathbf{i} = 0$ and

$$\boldsymbol{\mu}_t = \boldsymbol{\Phi}\boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t, \qquad Var(\boldsymbol{\eta}_t) = \boldsymbol{\Sigma}_\eta \tag{11}$$

or

$$\Delta\boldsymbol{\mu}_t = (\boldsymbol{\Phi} - \mathbf{I})\boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t.$$

Each row of $\boldsymbol{\Phi}$ sums to unity, $\boldsymbol{\Phi}\mathbf{i} = \mathbf{i}$. Thus setting λ to one in $(\boldsymbol{\Phi} - \lambda\mathbf{I})\mathbf{i} = \mathbf{0}$, shows that $\boldsymbol{\Phi}$ has an eigenvalue of one with a corresponding eigenvector consisting of ones. The other roots of $\boldsymbol{\Phi}$ are obtained by solving $|\boldsymbol{\Phi} - \lambda\mathbf{I}| = 0$; they should have modulus less than one for convergence.

If we write

$$\overline{\boldsymbol{\phi}}'\Delta\boldsymbol{\mu}_t = \overline{\boldsymbol{\phi}}'(\boldsymbol{\Phi} - \mathbf{I})\boldsymbol{\mu}_{t-1} + \overline{\boldsymbol{\phi}}'\boldsymbol{\eta}_t$$

it is clear that the $N \times 1$ vector of weights, $\overline{\boldsymbol{\phi}}$, which gives a random walk must be such that $\overline{\boldsymbol{\phi}}'(\boldsymbol{\Phi} - \mathbf{I}) = \mathbf{0}'$. Since the roots of $\boldsymbol{\Phi}'$ are the same as those of $\boldsymbol{\Phi}$, it follows from writing $(\boldsymbol{\Phi}'-\mathbf{I})\overline{\boldsymbol{\phi}} = \mathbf{0}$ that $\overline{\boldsymbol{\phi}}$ is the eigenvector of $\boldsymbol{\Phi}'$ corresponding to its unit root. This random walk, $\overline{\mu}_{\phi t} = \overline{\boldsymbol{\phi}}'\boldsymbol{\mu}_t$, *is a common trend in the sense that it yields the common growth path to which all the economies converge*. This is because $\lim_{j\to\infty} \boldsymbol{\Phi}^j = \mathbf{i}\overline{\boldsymbol{\phi}}'$; the proof follows along the same lines as that for a well-known result on ergodic Markov chains as given, for example, in Hamilton (1994, p681). The common trend for the observations is a random walk with drift, β, and with $\boldsymbol{\alpha}'\overline{\boldsymbol{\phi}} = 0$ each element of $\boldsymbol{\alpha}$ is a deviation from the common trend.

The *homogeneous* model has $\boldsymbol{\Phi} = \phi\mathbf{I} + (1-\phi)\mathbf{i}\overline{\boldsymbol{\phi}}'$, where $\mathbf{i}$ is an $N \times 1$ vector of ones, ϕ is a scalar convergence parameter and $\overline{\boldsymbol{\phi}}$ is an $N \times 1$ vector of parameters with the property that $\overline{\boldsymbol{\phi}}'\mathbf{i} = 1$. It is straightforward to confirm that $\overline{\boldsymbol{\phi}}$ is the eigenvector of $\boldsymbol{\Phi}'$ corresponding to the unit root. The convergence parameter and the elements of $\overline{\boldsymbol{\phi}}$, denoted $\overline{\phi}_i, i = 1, ..., N$ are estimated by maximum likelihood in the usual way by using the Kalman filter; the trend components are initialised with a diffuse prior. It is assumed that $|\phi| \leq 1$, with $\phi = 1$ indicating no convergence. The $\overline{\phi}_i$ parameters are constrained to lie between zero and one and to sum to one by maximising the log-likelihood with respect to N unconstrained parameters, a_i, that are linked to the $\overline{\phi}_i's$ by the equations $\overline{\phi}_i = a_i^2/\sum a_i^2, i = 1, .., N$.

Each trend can be decomposed into the common trend and a convergence component. The vector of convergence components is

$$\begin{aligned} \boldsymbol{\mu}_t^\dagger &= \boldsymbol{\mu}_t - \mathbf{i}\overline{\mu}_{\phi t} \\ &= \boldsymbol{\Phi}\boldsymbol{\mu}_{t-1} - \mathbf{i}\overline{\mu}_{\phi t-1} + \boldsymbol{\eta}_t - \mathbf{i}\overline{\eta}_{\phi t} = \phi(\mathbf{I} - \mathbf{i}\overline{\boldsymbol{\phi}}')\boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t - \mathbf{i}\overline{\eta}_{\phi t} \end{aligned}$$

so

$$\boldsymbol{\mu}_t^\dagger = \phi\boldsymbol{\mu}_{t-1}^\dagger + \boldsymbol{\eta}_t^\dagger, \qquad t = 1, ..., T.$$

where $\boldsymbol{\eta}_t^\dagger = \boldsymbol{\eta}_t - \mathbf{i}\overline{\eta}_{\phi t}$. Writing

$$\Delta\boldsymbol{\mu}_t^\dagger = (\phi - 1)\boldsymbol{\mu}_{t-1}^\dagger + \boldsymbol{\eta}_t^\dagger, \tag{12}$$

shows that each relative growth rate depends on the gap between the series in question and the common trend. Substituting into (10) gives

$$\mathbf{y}_t = \boldsymbol{\alpha}+\beta\mathbf{i}t + \mathbf{i}\overline{\mu}_{\phi t} + \boldsymbol{\mu}_t^{\dagger} + \boldsymbol{\psi}_t+\boldsymbol{\varepsilon}_t, \quad t = 1, ..., T$$

Once convergence has taken place, the model is of the form (9), but with an additional stationary component $\boldsymbol{\mu}_t^{\dagger}$.

The smooth convergence model is

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\mu}_t + \boldsymbol{\psi}_t+\boldsymbol{\varepsilon}_t, \quad t = 1, ..., T \tag{13}$$

with $\boldsymbol{\alpha}'\mathbf{i} = 0$ and

$$\begin{aligned} \boldsymbol{\mu}_t &= \boldsymbol{\Phi}\boldsymbol{\mu}_{t-1} + \boldsymbol{\beta}_{t-1} \\ \boldsymbol{\beta}_t &= \boldsymbol{\Phi}\boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t \end{aligned}$$

The forecasts converge to those of a smooth common trend, but in doing so they may exhibit temporary divergence; see Harvey and Carvalho (2002).

In scalar notation the homogeneous model can be expressed in terms of the common trend, $\overline{\mu}_{\phi,t}$, and convergence processes, $\mu_{it}^{\dagger} = \mu_{it} - \overline{\mu}_{\phi,t}, i = 1, ..., N$, by writing

$$y_{it} = \alpha_i + \overline{\mu}_{\phi,t} + \mu_{it}^{\dagger} + \psi_{it} + \varepsilon_{it}, \quad i = 1, ..., N, \tag{14}$$

where $\Sigma\alpha_i = 0$ and

$$\mu_{it}^{\dagger} = \phi\mu_{i,t-1}^{\dagger} + \beta_{it}^{\dagger}, \quad i = 1, ..., N, \qquad |\phi| < 1 \tag{15}$$

$$\beta_{it}^{\dagger} = \phi\beta_{i,t-1}^{\dagger} + \eta_{it}^{\dagger}, \tag{16}$$

where the initial conditions, $\mu_{i0}^{\dagger}, i = 1, .., N-1$, are fixed and

$$\begin{aligned} \overline{\mu}_{\phi t} &= \phi\overline{\mu}_{\phi t-1} + \overline{\beta}_{\phi t-1} \\ \overline{\beta}_{\phi t} &= \phi\overline{\beta}_{\phi,t-1} + \overline{\eta}_{\phi t}. \end{aligned}$$

Note that the $N-th$ series (or indeed any series) can be constructed from the others as

$$\mu_{Nt}^{\dagger} = -\overline{\phi}_N^{-1}\sum_{i=1}^{N-1}\overline{\phi}_i\mu_{it}^{\dagger} \quad with \quad \overline{\phi}_N = 1 - \sum_{i=1}^{N-1}\overline{\phi}_i.$$

and similarly for $\beta_{Nt}^{\dagger}$.

4 Stylised facts on trends and convergence in Euro-zone Countries

In this section we use the multivariate model of sub-section 2.2 to display the stylised facts concerning trends and convergence in real per capita incomes in eleven Euro-zone countries: Austria (AU), Belgium (BE), Finland (FI), France (FR), Germany (GE), Greece (GR), Ireland (IR), Italy (IT), Netherlands (NE) Portugal (PO) and Spain (SP). Out of the present membership of the Euro-area, only Luxembourg is not included as no data are available.

4.1 Data and estimation

Data were obtained from the GGDC Total Economy Database, 2002, at the University of Groningen and the Conference Board [4]. All series are expressed in 1990 US dollars converted at ‘Geary-Khamis’ purchasing power parities and log-transformed. Figure 1 shows annual observations from 1950 to 1997. The German data refers to West Germany. After 1997 data is only available for Germany as a whole and this is why we concentrated our analysis on the period up to that point.

Estimation of all the new model was done using program routines written in the `OX 3.0` language (Doornik, 1999) with use being made of the SSfPack package for state space algorithms of Koopman, Shephard and Doornik (1999). Some of the more standard models were estimated using the STAMP package of Koopman et al (2000). All parameters were estimated by maximum likelihood as described in Section 2 and variances are reported multiplied by 10^5. The estimated covariance matrices are reported by showing the variances on the main diagonal while the entries above contain the cross-correlations. The components shown in the figures are extracted by the state-space smoothing algorithm. Comparisons with US regions refer to the results in Carvalho and Harvey (2002).

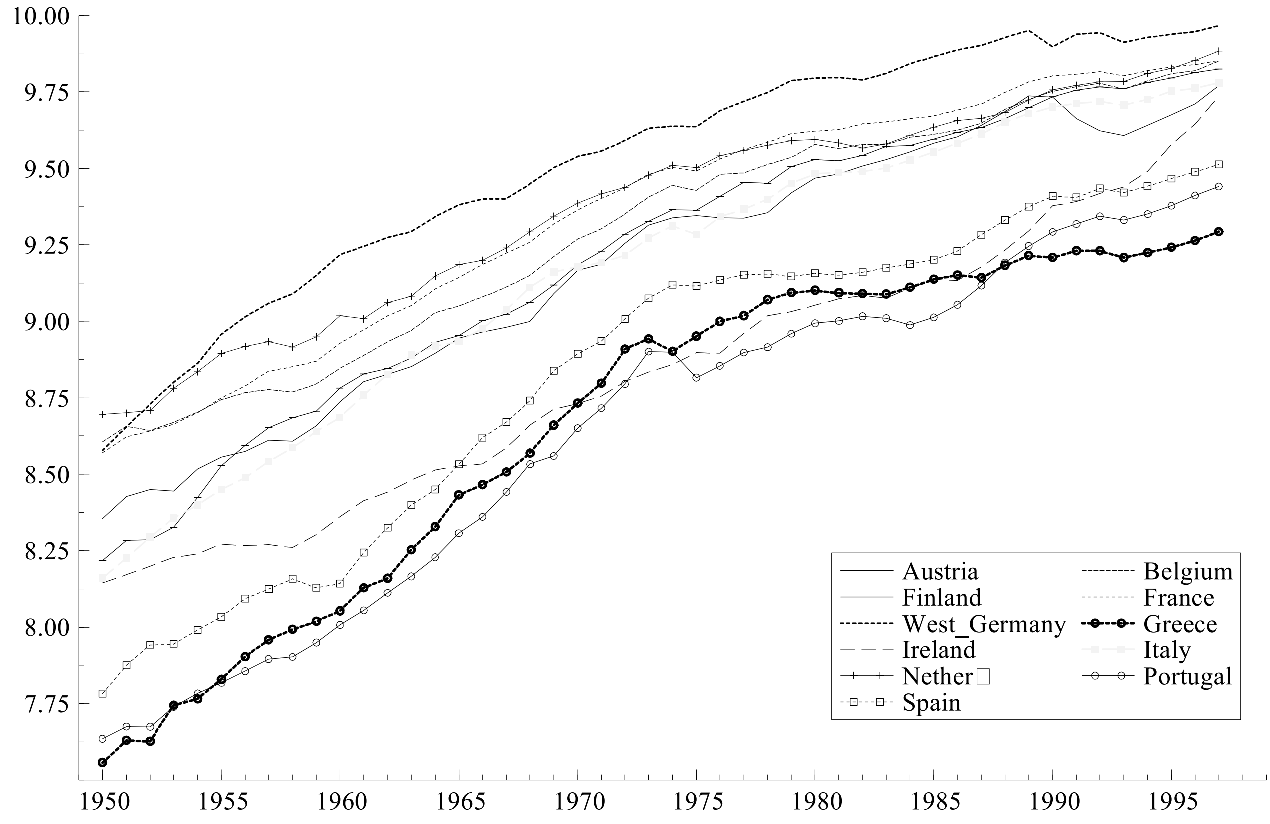

Figure 1: Annual series (1950-1997) for Euro-area countries

[4]For full description of the dataset refer to http://www.eco.rug.nl/ggdc

4.2 Trend growth and income inequality

Figure 2 shows the trends, $\widetilde{\mu}_{it|T}$, obtained by fitting a simple multivariate smooth trend plus cycle model. These slowly changing trends indicate the long-run movements from which we can start to assess any tendencies towards convergence or divergence.

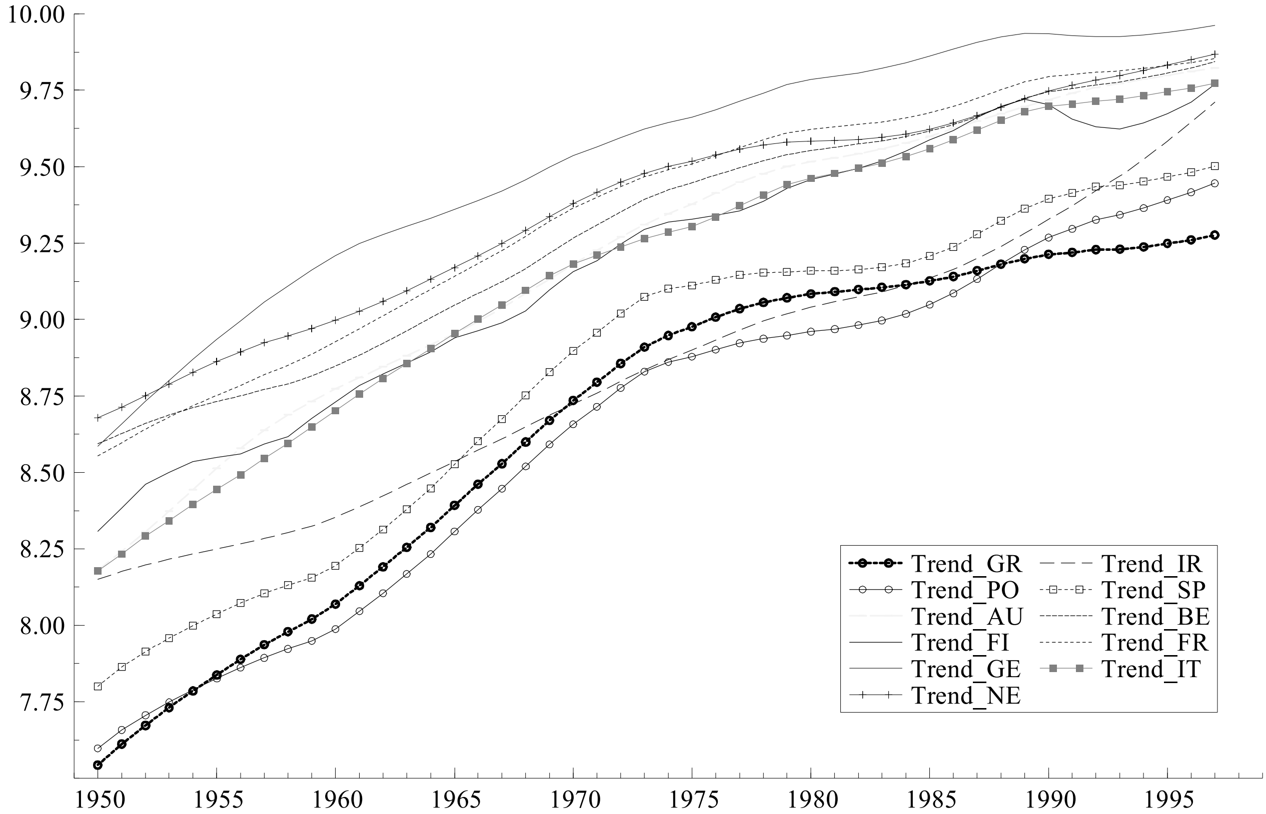

Figure 2: Smooth trends for Euro-zone countries

The first point to note is the existence of subgroups, or clusters, within the Euro-zone: a high-income group, consisting of traditional core countries (GR, FR, BE, NE), together with some peripheral countries (IT, AU, FI), and a low income group made up of poorer peripheral countries (PT, SP, GR). The outlier is Ireland, which cannot be labeled with certainty as a member of either group given its growth dynamics. Otherwise, this subgrouping is well defined and remarkably stable throughout the second half of the twentieth century. The three poorest and the seven richest countries in 1950 still had that status at the end of the nineties.

The second point to note is that the extracted trends clearly show the three distinct epochs of economic growth: 1950-1972; 1973-1979 and 1980-1997. These standard subdivisions in post-war European economic history correspond to the 'Golden Age', the shocks of the 1970s and the ensuing period of stabilisation and restructuring; see Crafts and Toniolo (1996). Table 1 below documents the average trend growth rates for each of the Euro-zone countries and for each subgoups.

Table 1: Average trend growth rates for Euro-zone countries

	1950-1997	1950-1972	1973-1979	1980-1997
Austria	3.57%	5.10%	3.35%	1.80%
Belgium	2.70%	3.51%	2.71%	1.70%
Finland	3.19%	4.38%	2.65%	1.94%
France	2.81%	4.09%	2.54%	1.36%
Germany	3.00%	4.71%	2.51%	1.09%
Greece	3.79%	6.15%	3.12%	1.15%
Ireland	3.38%	2.99%	3.21%	3.93%
Italy	3.47%	4.94%	2.95%	1.86%
Netherlands	2.57%	3.57%	1.90%	1.62%
Portugal	4.03%	5.51%	2.49%	2.81%
Spain	3.72%	5.72%	1.99%	1.95%
7 richest, average	3.04%	4.33%	2.66%	1.63%
3 poorest, average	3.85%	5.79%	2.53%	1.97%

Taking the information in Figure 2 and Table 1 together, it is clear that: 1) the poorest subgroup has, over the entire period, a higher trend growth rate; 2) this higher average trend growth rate is a product of the strong catching-up process during the Golden Age[5].

These two facts largely explain the findings in cross-sectional studies, such as Sala-i-Martin (1996), supporting β-convergence in Europe throughout the post-war period and its absence from the 1970s onwards. Table 1 also sheds some light on (limited) mobility in income per-capita ranking within each group.

Thus, for example Austria's and Portugal's superior growth performance leads to their overtaking of Finland and Greece respectively. On the other hand, Ireland's idiosyncratic behaviour is singled out: it is both the only country to miss out on the Golden Age and the best achiever over the 1980-1997 period.

Further insights on the evolution of income inequality across the Euro-zone can be gained by the plotting the cross-sectional standard deviation of the smoothed trend components, $SD(\widetilde{\mu}_{it|T})$. This is done in Figure 3, where we also plot the analogous series for the 8 US Census regions, as well as those of subgroups both within Europe and US.

Some interesting facts arise from Figure 3. The first is that overall Euro-zone trend dispersion (solid line) has in the last half-century. As was clear from the analysis of Table 1, this is mainly the result of the Golden Age period. One important factor is trade liberalisation process, set in motion in the 1960s both within the EEC and EFTA and between the two organisations; see Ben-David (1993). However the fact remains that, by the end of the nineties, inequality across the Euro-zone was still above that of the US regions (dotted line) in the 1950s. The analysis of inequality within the two subgroups in Europe allows us to qualify the last statement.

[5]Calculations by the authors show that this catching-up process is largely limited to the 60s. During the 50s the high income group actually grows faster due to the high growth rates of Austria, Germany and Italy, i.e. reconstruction efforts still appear to matter in explaining the 1950s relative performance.

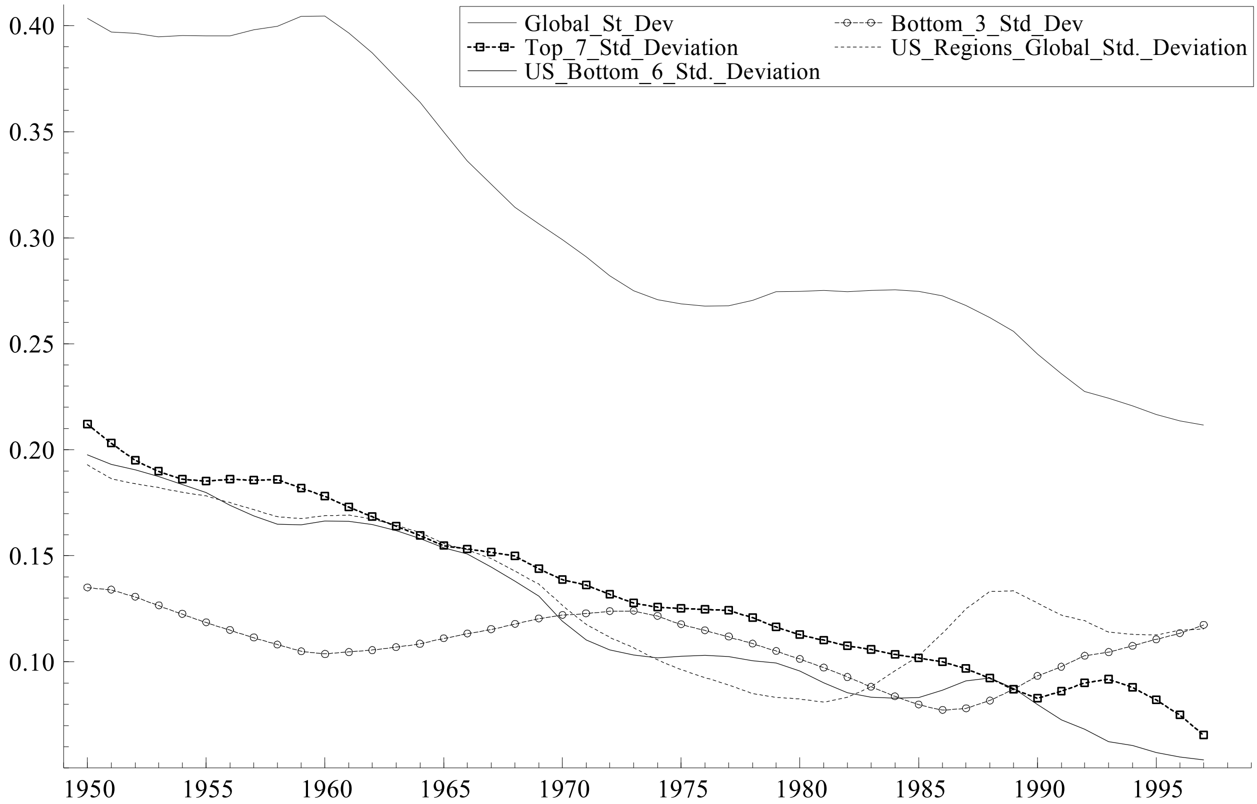

Figure 3: Evolution of global and subgroup cross-sectional standard deviations in Euro-zone and US

From Figure 3, it is clear that inequality within the groups is much lower than it is overall. In particular, inequality within the high-income group (box line) has been falling steadily and since the mid-eighties it has been below that of the US regions. Within the low-income group, inequality in the 1950s was relatively low and no downward trend can be observed in the data. Nevertheless, it is still the case that relative to the US benchmark, inequality within both groups of countries is not very different.

4.3 Convergence and multiple steady states

The preliminary investigation of stylised facts on trend growth in within current Euro-zone countries indicates the existence of two convergence clubs[6], plus Ireland. If this is the case, convergence tests and models based on the assumption of a common growth path will be inadequate and could lead to misleading inferences being drawn. What we need to do is to investigate the case for absolute convergence within the two groups and relative convergence between the groups.

[6]Durlauf and Johnson (1995) identify four different regimes (determined by initial conditions) in the Summers-Heston dataset. They find that these regimes are associated with different aggregate production functions and hence, the assumption of a common long-run growth path is untenable. Their partition of the eleven Euro-zone countries is the same as the one noted in the previous subsection. Thus, Greece, Portugal, Spain (and Ireland) are classified as intermediate-output countries with high literacy rates while the remaining seven are included in the high-output group.

5 Fitting convergence models

Based on the analysis of the previous section, we proceed by fitting convergence models, as described in section 3, to the rich and poor subgroups rather than to the Euro-zone as a whole.

Recall that this specification not only allows us to separate trends and cycles but also separates out the long-run balanced growth path from the transitional (converging) national dynamics, thus permitting a characterisation of convergence stylised facts.

The results reported are for the homogeneous model with smooth convergence, (13), and absolute convergence, that is $\alpha_i = 0, i = 1, ..., N$. We analyse each sub-group in turn.

5.1 High-Income Group

Figure 4 below displays the estimated common trend, $\overline{\mu}_{\phi,t}$, together with the estimated trends for each country while Figure 5 shows the smoothed estimates of the convergence components, $\mu^{\dagger}_{i,t}$, for the seven Euro-zone high-income countries [AU, BE, FI, FR, GE, IT, NE].

The convergence parameter, ϕ, was estimated[7] as 0.938 while the common trend weights, $\overline{\phi}_i$, are given by:

$\overline{\phi}_{AU}$	$\overline{\phi}_{BE}$	$\overline{\phi}_{FI}$	$\overline{\phi}_{FR}$	$\overline{\phi}_{GE}$	$\overline{\phi}_{IT}$	$\overline{\phi}_{NE}$
0.01	0.06	0.45	0.02	0.03	0.40	0.03

The within-group convergence is shown in Figures 4 and 5, with all seven countries significantly narrowing their gap towards the common balanced growth path.

The large weights assigned to Finland and Italy in constructing the common trend means that the future growth path depends primarily on extrapolating the trends in these two countries. In other words these two relatively poorer countries within the high-income group act as benchmarks to which all the other high-income countries converge.

This may seem surprising at first. Why doesn't Germany dominate, for example? A glance at Figure 4 gives the answer. The German growth rate has been gradually slowing down, particularly after re-unification, and at the end of the sample the trend is almost flat.

If we were to extrapolate there would be almost no growth.

[7] The convergence parameter in a first-order error correction model, (12), will typically be higher.

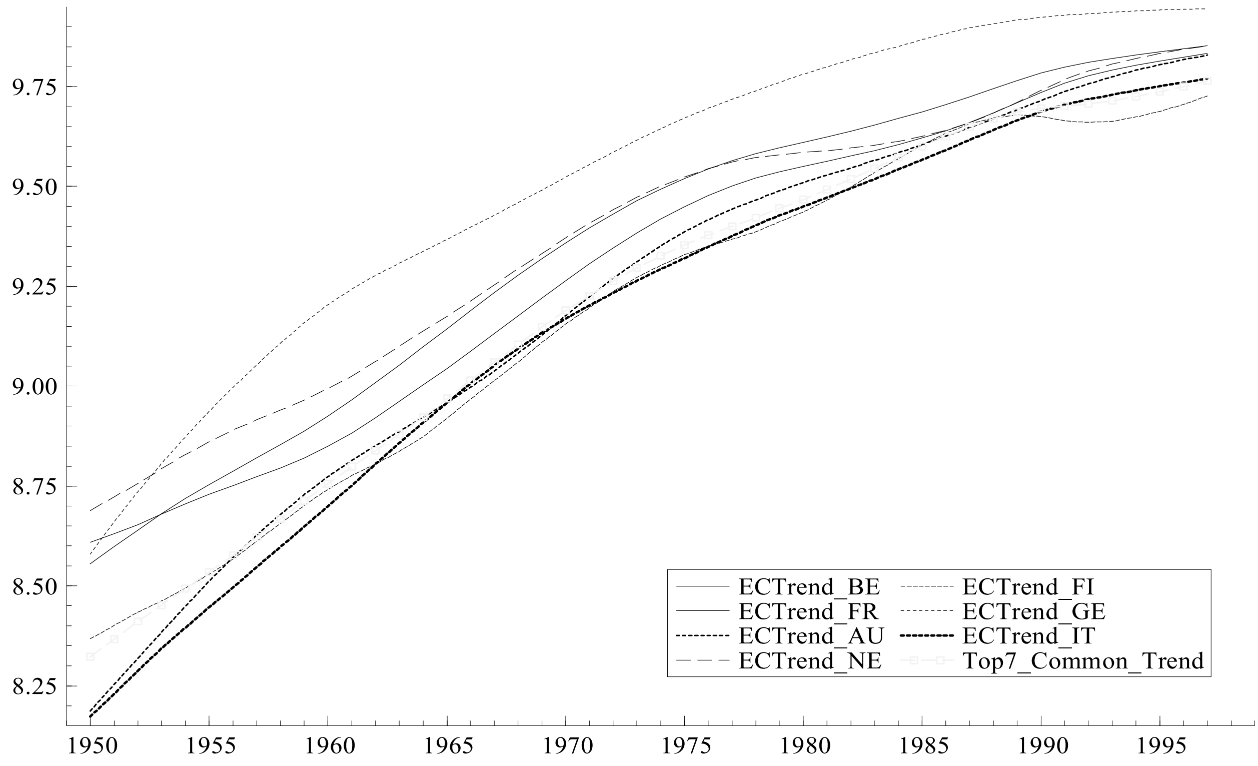

Figure 4: Converging trends and balanced growth path for the high income group

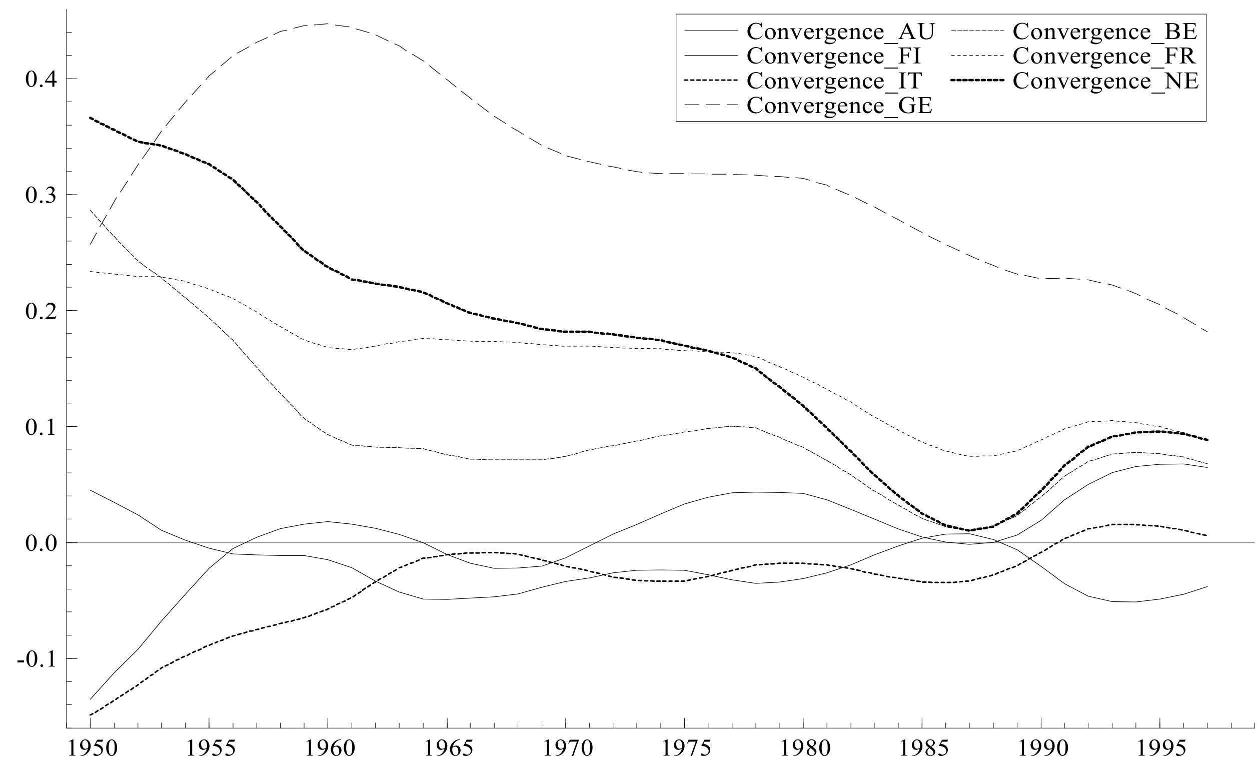

Figure 5: Convergence components for high-income group

The convergence components are not monotonic over time and so the fact that the model allows a degree of divergence is important. During the 1950s and early 60s at least two distinct type of dynamics should be singled out. One is dominated by the strong effects of reconstruction in Austria, Italy and Germany and the other by relatively slower adjustment in France, Belgium and the Netherlands; further details can be found in Crafts and Toniolo (1996) and the references therein. In contrast the late 1960s and 1970s appear to bring a halt in the convergence process across the entire high-income group. This process is resumed in the 1980s before and again slowing down in the 1990s.

Further insights on cross national correlations can be gained by analysing the variance-covariance matrix of the converging trends, $\widetilde{\boldsymbol{\Sigma}}_{\zeta}$. Two subgroups, with relatively high, positive within-correlations and relatively low (or negative) across-correlations appear to emerge. On the one hand a group formed by Belgium, France, Netherlands and, to a smaller extent, Italy and on the other hand, a group formed by Austria, Germany and possibly Finland.

$$\begin{bmatrix} 1.80 & 0.31 & 0.15 & -0.03 & 0.49 & -0.48 & 0.48 \\ & 1.22 & -0.19 & 0.83 & -0.63 & 0.09 & 0.94 \\ & & 8.15 & 0.18 & 0.57 & 0.03 & -0.26 \\ & & & .86 & -0.60 & 0.48 & 0.70 \\ & & & & .48 & -0.36 & -0.47 \\ & & & & & 1.06 & 0.15 \\ & & & & & & 2.23 \end{bmatrix} \begin{matrix} AU \\ BE \\ FI \\ FR \\ GE \\ IT \\ NE \end{matrix}$$

The forecasts of the convergence components will tend to zero as the lead time goes to infinity. The second-order ECM allows some divergence before convergence eventually takes place; see the model fitted to the US and Japan in Harvey and Carvalho (2002). In Figure 5 all the convergence components are pointing in the direction of zero and a straightforward extrapolation suggests that the economies will be close to convergence after about ten years.

5.2 Low-Income Group

We now turn to the converging dynamics within the low income group [GR, PO, SP]. Ireland is excluded due to its idiosyncratic behaviour: it is difficult to classify it as a low income country given its position both at the beginning and end of the sample[8]. Figures 6 and 7 below present the common trend and convergence components respectively.

The convergence parameter was estimated as $\phi = 0.937$, and the common trend weights, $\overline{\phi}_i$ given by:

$\overline{\phi}_{GR}$	$\overline{\phi}_{PO}$	$\overline{\phi}_{SP}$
0.48	0.36	0.17

[8]Taking the Irish performance after the end of our sample (1997) only serves to confirm this assumption as the gap between Ireland and the three low income countries grows even larger.

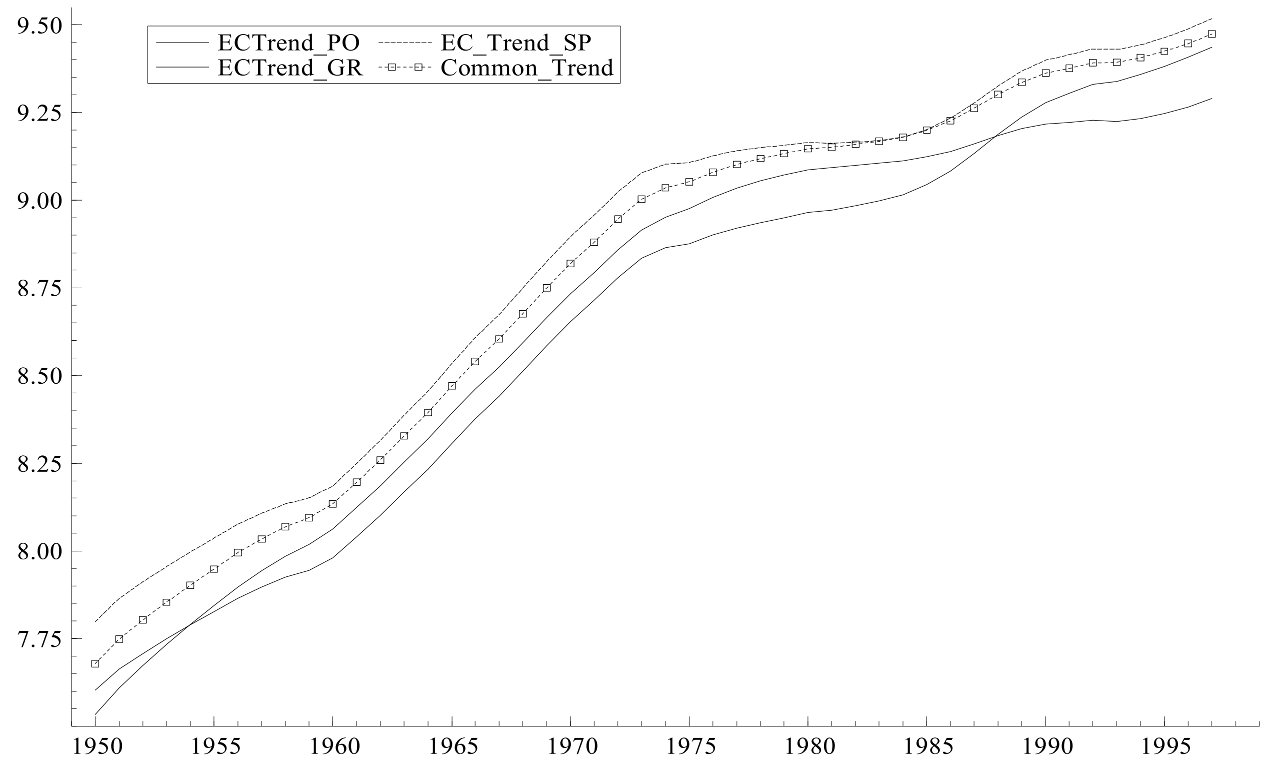

Figure 6: Common trend and converging trends for low-income group

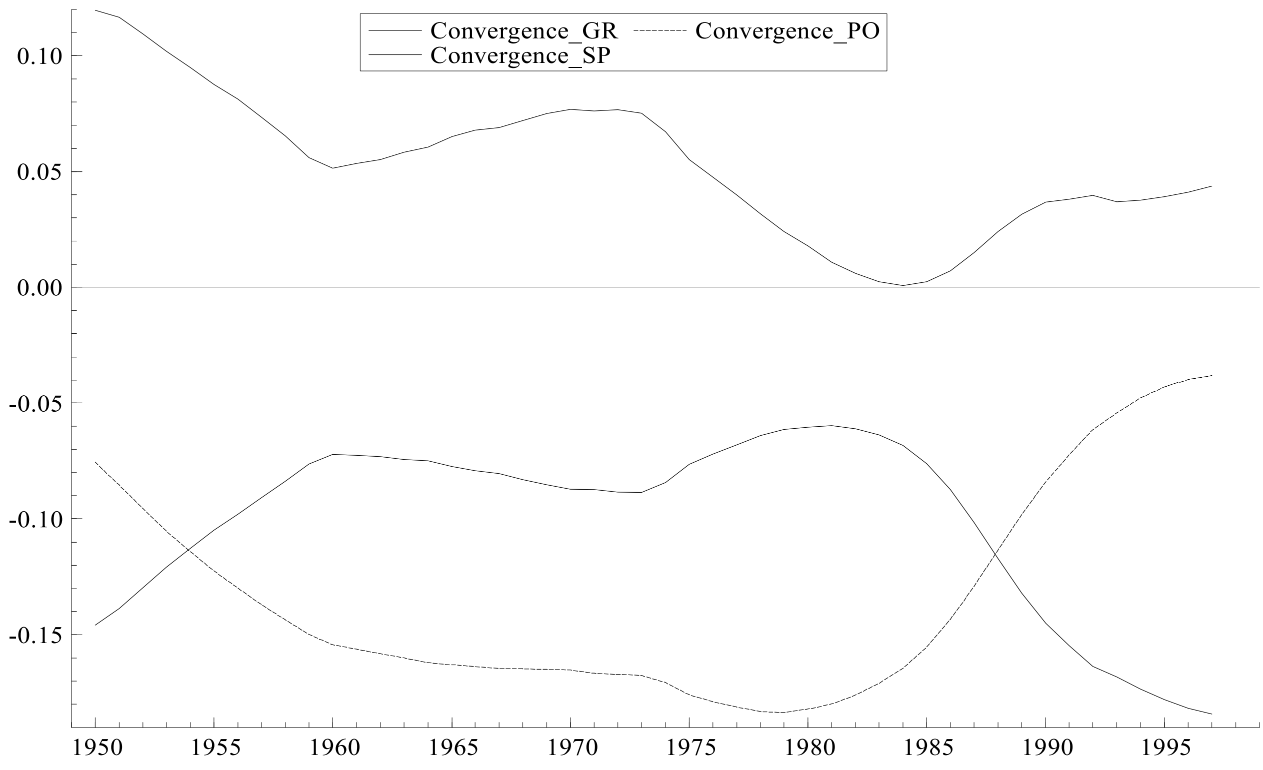

Figure 7: Convergence components for low income group

The sign of the convergence parameter is consistent with within-group convergence and similar in magnitude to the high-income group. The cross-correlations in the covariance matrix of the converging trends, $\widetilde{\mathbf{\Sigma}}_{\zeta}$ are shown below.

$$\left[\begin{array}{ccc} 10.0 & 0.89 & 0.95 \\ & 17.5 & 0.99 \\ & & 22.8 \end{array}\right] \begin{array}{c} GR \\ PO \\ SP \end{array}.$$

All the pairwise correlations are greater than the BE-FR correlation, the largest in the high-income group. The magnitude of the convergence components is smaller and more stable for these three low income countries. However, Figure 7 illustrates the point that a relatively stable cross-sectional standard deviation does not necessarily imply that there are no intra-distribution dynamics; see Quah (1996) and Durlauf and Quah (1999). Thus, Greece and Portugal alternate their ranking over the sample period: Greece starts off as the poorest economy, rapidly overtakes Portugal, but is subsequently overtaken by Portugal in the 1980s. Indeed the main conclusion to be drawn from the plots of the convergence components is that since that mid-eighties the gap between Spain and Portugal has narrowed considerably, while Greece has diverged. This throws some doubt on the coherence of the group.

5.3 Divergence or relative convergence in the Euro-zone?

Having characterised the within-group convergence dynamics we now proceed with a between groups analysis. The key question is whether the common trends estimated for each group indicate relative convergence. In order to investigate this issue, a bivariate convergence model is fitted to the two common trends. Since we are dealing with extracted trends, $\widetilde{\mu}_{it}$, the aim is to perform the decomposition:

$$\widetilde{\mu}_{it} = \widetilde{\alpha}_i + \widetilde{\overline{\mu}}_{\phi,t} + \widetilde{\mu}^{\dagger}_{it}, \quad i = LI, HI \tag{17}$$

with $\overline{\mu}_{\phi,t}$ and $\mu^{\dagger}_{it}$ defined as in (14) and LI and HI denoting low income and high income, respectively.

When the model is fitted with no restrictions on ϕ_{LI} and ϕ_{HI}, the estimate of ϕ_{HI} is zero, indicating that the high-income trend can be treated as benchmark. Table 2 below shows the parameter estimates for the absolute and relative convergence versions of this benchmark model. It is convenient to standardise the intercepts in the relative convergence model by setting $\widetilde{\alpha}_{HI} = 0$; there is then absolute convergence when $\widetilde{\alpha}_{LI} = 0$.

Table 2: Estimates for bivariate convergence model fitted to smoothed trends

		Absolute		Relative	
	Hyperparameters	LI	HI	LI	HI
Convergence	$\sigma^2_{\zeta}(\times 10^5)$	6.85	.54	6.51	.54
	$\widetilde{\phi}$	0.939		0.916	
Gap	$\widetilde{\alpha}_{LI}$	$0(fixed)$		-0.337 (0.168)	
Fit	$\log L$	366.56		367.92	

Two main points should be noted from Table 2. First, the estimate of ϕ is less than one and so is consistent with convergence of the low-income trend towards the high income trend. Second, the estimate of α_{LI} is negative and its 't-statistic' is statistically significant[9] at the 5% level.

This implies that the trend forecasts of the low income countries are converging to a growth path, $\widetilde{\mu}_{HI,t} + \widetilde{\alpha}_{LI}$, that is parallel to that of the high income countries but lies below it.

Figure 8 shows a plot of $\widetilde{\mu}_{HI,t} + \widetilde{\alpha}$, labelled 'common trend', together with the two group common trends. The convergence component, $\widetilde{\mu}^{\dagger}_{LI,t} = \widetilde{\mu}_{LI,t} - (\widetilde{\mu}_{HI,t} + \alpha)$, is shown in Figure 9.

It can be seen that relative convergence had taken place by the early seventies. The gap appears to be permanent and the long-run implication is that each of the low income countries will have a per capita income that is almost 30% below that in the high income group.

Figure 8: Common trend and trends for low income and high income group

[9] The LR statistic is not significant however. Since a good deal of the variation has been removed by smoothing, tests of significance should only be regarded as giving a rough indication of the underlying situation.

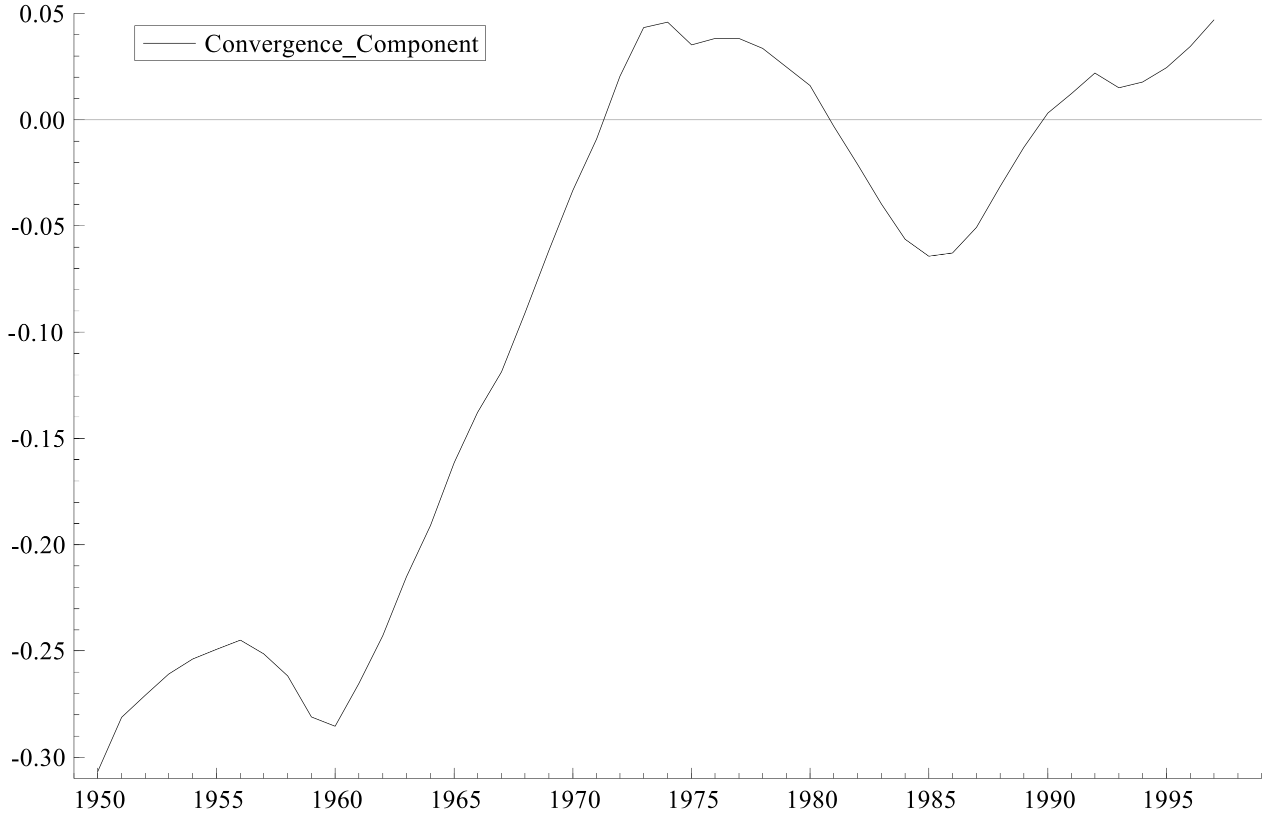

Figure 9: Convergence component for low income group

6 Cycles

The other aspect of convergence concerns cycles. In order to investigate this issue we looked at the five core economies of France, Germany, Belgium, Netherlands and Italy to see to what extent the business cycle movements had converged. All of these countries are in the rich group. We have excluded Austria and Finland as they are relatively recent entrants into the EC. Fitting a multivariate STM to the five series using the STAMP package of Koopman et al (2000) gave very similar results to fitting the convergence model reported in sub-section 5.1 and it is quite adequate for the purposes of extracting cycles. The model consisted of smooth trends, similar cycles and irregular components.

Focussing attention on the cycle, the estimates of the damping factor, ρ, and the period were found to be 0.87 and 7.86 years respectively. The variances and cross-correlations are shown below.

BE	15.1	.75	.79	.55	.28
FR		12.2	.78	.45	.34
GE			24.8	.72	.35
NE				25.6	.26
IT					31.6

In our study of US regions we performed a principal components analysis on the correlation matrix and showed that the biggest component accounted for over 90% of the variation. This was not surprising in view of the fact that the majority of pairwise correlations were greater that 0.9.

The issue here is somewhat different in that our interest is on the extent to which the cycles have come closer together. Figure 10 shows the smoothed estimates of the cycles. While the dispersion at the beginning is quite large, the cycles are almost perfectly co-ordinated by the end of the 1990s. Figure 11, which plots the standard deviation of the five cycles over time, tells the same story.

As part of the convergence process, the cycle cross-correlations may well have changed over time so the figures above have to be regarded as averages in some sense.

Estimating the model over the period starting in 1970 gave estimates of ρ and the period equal to 0.90 and 5.90 years respectively. As can be seen below, the cross-correlations are much bigger.

BE	29.8	.88	.64	.78	70
FR		19.1	.78	.86	.69
GE			9.5	.75	.55
NE				25.1	.53
IT					23.7

As it stands, the cycle model does a good job of picking up the convergence, but if it were to be used to generate converging cycles, it would have to incorporate a mechanism that allowed the cross-correlations to tend gradually towards unity.

The same model was also fitted to three poorer countries, Greece, Portugal and Spain. The damping factor and the period were found to be 0.81 and 7.29 years, while the cycle variances and cross-correlations are

GR	22.3	.43	.20
PO		67.8	.05
SP			23.7

The relatively low value of ρ indicates that the cycles are not particularly well-defined. This is especially true of Spain. The low correlations between Spain and the other two countries indicates little coherence. Portugal's cycles are quite pronounced, with the variance being three times what it is in most countries.

The cycles are shown in Figure 11. Despite the low coherence in the period as a whole, it can be seen that they are converging towards the end of the period, with a pattern not dissimilar to that in the core group.

Finally a cycle was extracted for Ireland by modelling it jointly with Germany. This gave a moderately high correlation of 0.79 but with Ireland having a much larger variance of 49.6 ($\times 10^{-5}$). The extracted cycles are shown in Figure 12. The most striking feature is the higher amplitude of the Irish cycle, even in recent years.

7 Testing re-visited

As was noted earlier, standard unit root tests give a confused message with regard to convergence. This section examines some of the issues in testing for convergence in the light of the descriptive results of the last two sections. In doing so we draw on the recent investigation by Harvey and Bates (2003).

The first point to note is that the standard augmented Dickey-Fuller (ADF) test, with a constant included, has very low power. This can be illustrated by applying the test to the series on The Netherlands minus Italy. A graph shows very clear convergence. In 1950 the difference (in logs) between the two series was 0.534, corresponding to The Netherlands having a per capita income 1.7 times that of Italy, while in 1997, the gap was 0.105, corresponding to a ratio of 1.11. Yet the ADF test fails to reject. For example, with one lagged first difference[10], the $t-$statistic is -2.14, as against a 10% critical value of -2.60.

This is despite the fact, that, an initial value away from the mean actually helps to increase power. If the constant is dropped from the ADF regression, as is appropriate for a test of absolute convergence, then, with one lag, the $t-$statistic is -2.52, as against a 5% critical value of -1.96.

Thus dropping the constant leads to a rejection of the null of no convergence at the 5% level of significance. The t-statistic is not quite significant at the 1% level, but if the information in all five core countries is pooled by forming a set of four contrasts with Italy, the multivariate homogeneous Dickey-Fuller t-test, studied in Harvey and Bates (2003), shows a very strong rejection[11]. Taking the observations from 1960 (since there is some divergence before 1960, although not between The Netherlands and Italy); the $t-$statistic, with one lag, is -3.04 while the 1% critical value (for $T = 100$) is -2.58.

If all seven rich countries are considered, the rejection is even stronger as the $t-$statistic is -4.27. On the other hand if the constant is included, even the multivariate tests are unable to reject at the 10% level[12]. Note that because the test takes account of cross-correlations between the series, it is invariant to the choice of benchmark.

Testing in the low income group gives a less clear picture, as one might expect from the earlier analysis. Pairwise tests are unable to reject the null of no convergence, even with the constant excluded, and pooling the data does not help[13].

[10] The result is relatively insensitive to the number of lags. For example with three lags $t = -2.41$. (These additional lags are not statistically significant at the 10% level.)

[11] A LR test of the null hypothesis that the four contrasts are nonstationary is unable to reject at the 5% level of significance; whether or not intercepts are included makes no difference to the conclusion. Such a result is consistent with the simulations in Harvey and Bates (2003), which indicate that the LR test can have very low power relative to the multivariate homogeneous Dickey-Fuller t-test.

[12] For 5 and 7 countries the t-statistics were -2.26 and -2.25 respectively. The 10% critical values - from table 3 of Harvey and Bates(2003) are -3.73 and -4.26 respectively.

[13] The t–statistic, with no constant, for Greece-Spain with one lag is only -0.069, while Portugal-Spain is -0.924 and -1.070 if a constant is added.

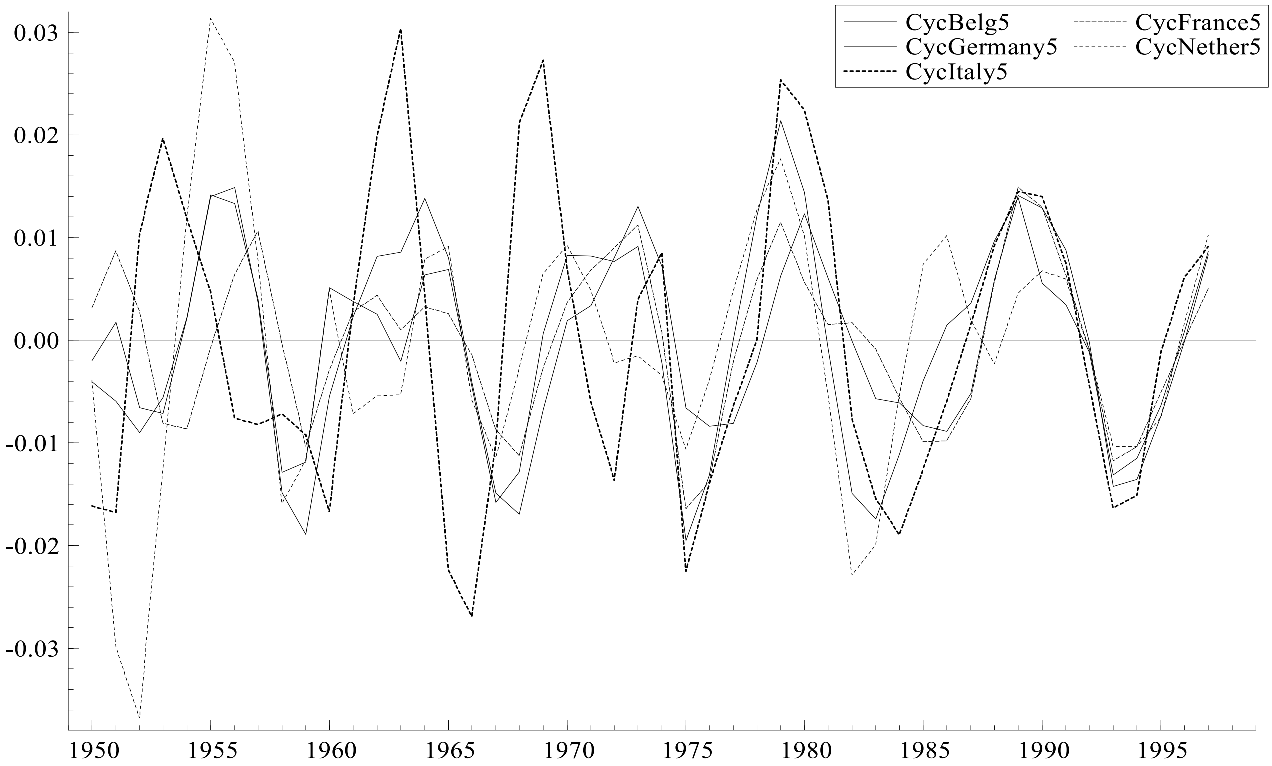

Figure 10: Cycles extracted from five core countries

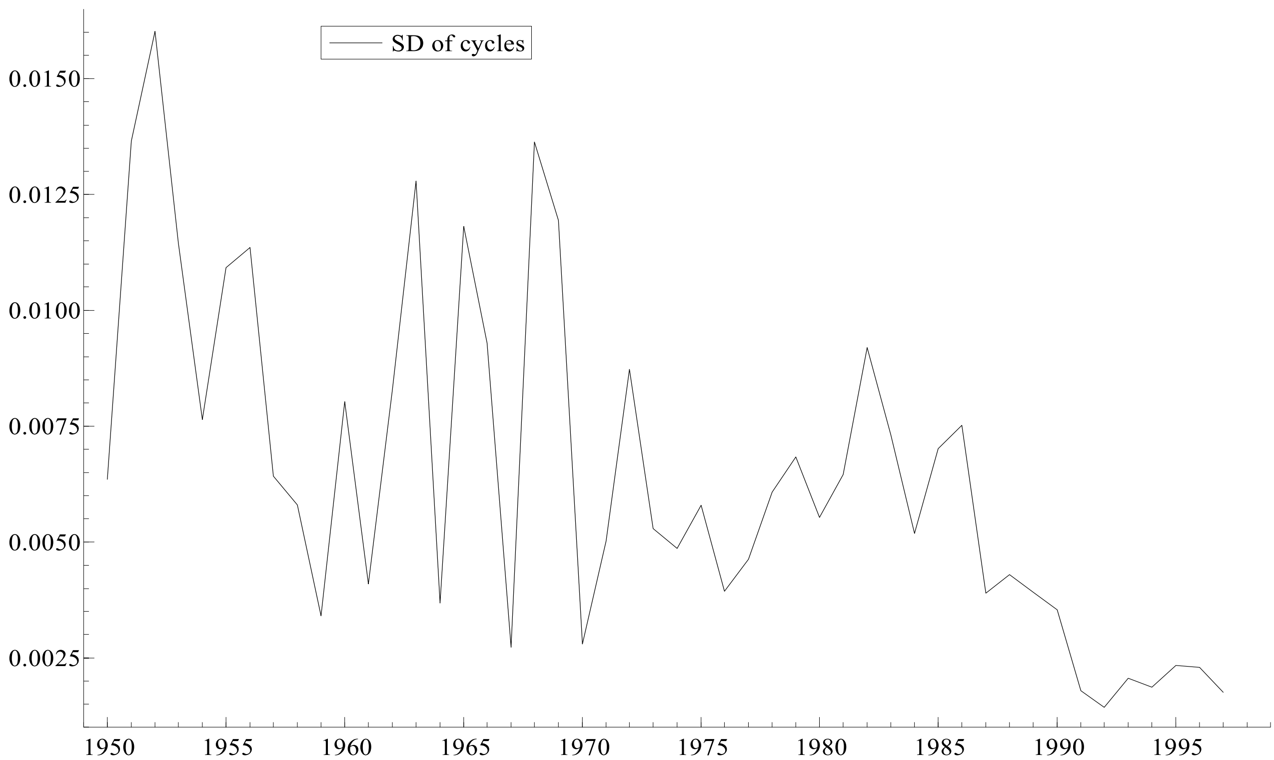

Figure 11: Standard deviation of cycles extracted from five core countries

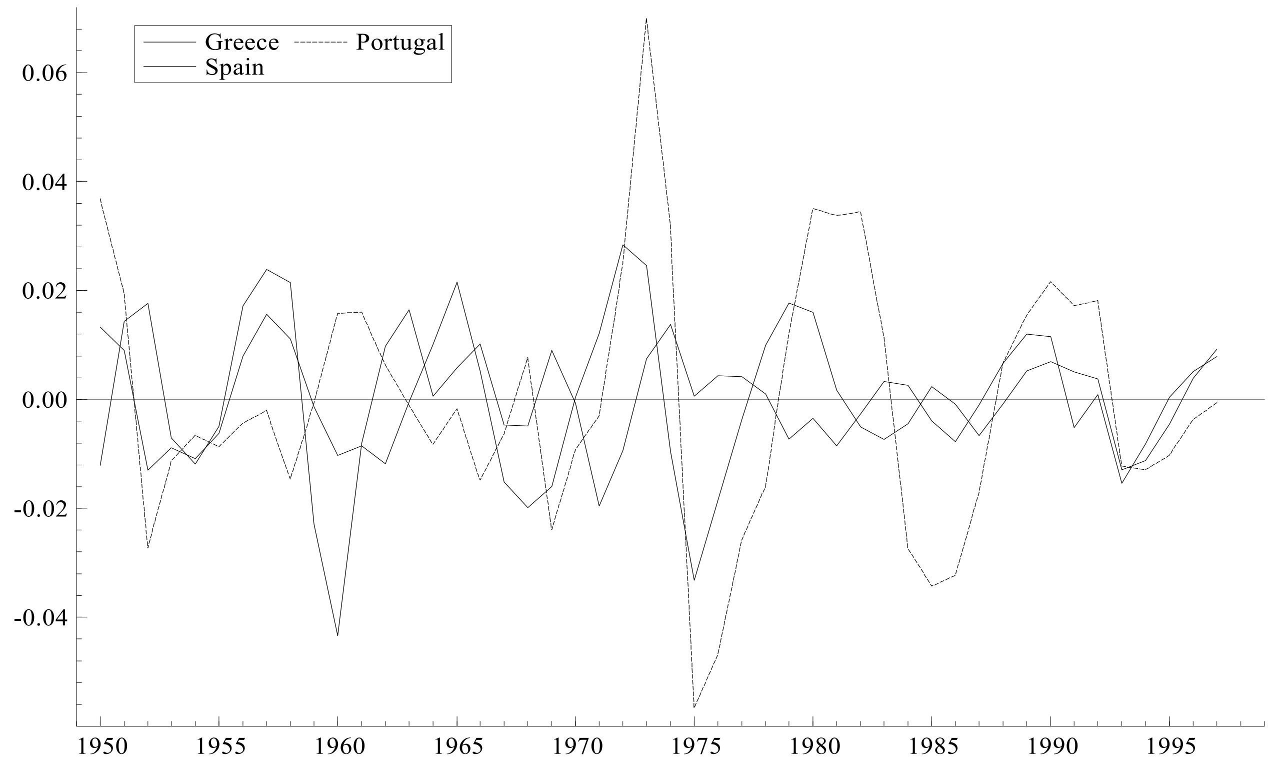

Figure 12: Cycles for Greece, Portugal and Spain

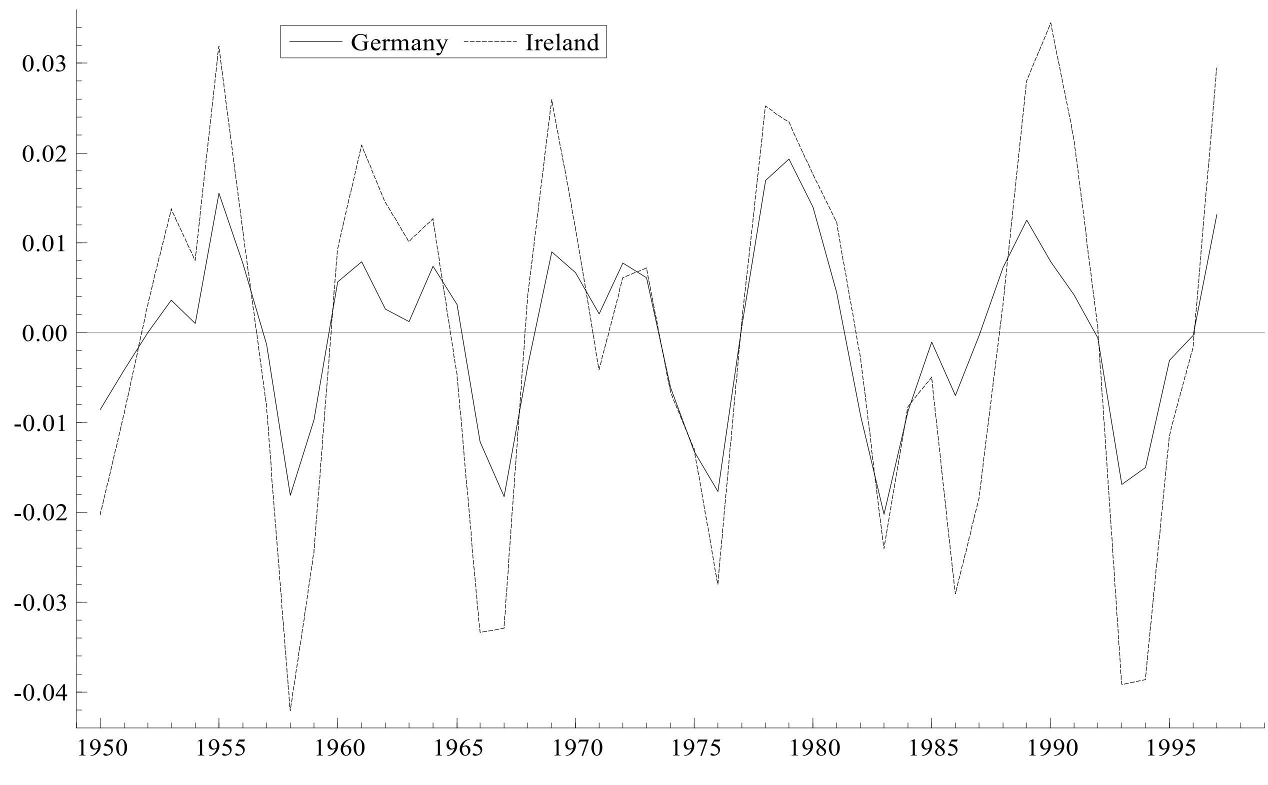

Figure 13: Cycles for Ireland and Germany

The trace likelihood ratio test of Johansen (1988) may be used, as in Bernard and Durlauf (1995), to try to detect the number of co-integrating relationships and hence the number of clubs. However, the test is likely to have low power and hence to indicate too many clubs. For example with the five core countries, again from 1960, the test, based on a model with an unrestricted constant[14] and two (p) lags, cannot reject the null hypothesis of five common trends, against the alternative that there are fewer, at the 5% level of significance when the degrees of freedom correction is used. In other words it indicates that there is no co-integration and no convergence clubs (with more than one member). With no degrees of freedom correction, four trends cannot be rejected. The results, obtained from the PcFIML program of Doornik and Hendry (2000), are shown in the Table below; R is the number of co-integrating vectors so $5 - R$ is the number of (common) trends. The max test, not shown in the table, could not reject five trends even without the degrees of freedom correction.

Table 3: Trace LR statistic for co-integration with $(T - Np)$ and without (T) d.o.f. correction

R	T	$T - Np$	5% cv
0	70.57*	52.00	68.5
1	42.79	1.53	47.2
2	19.74	14.54	29.7
3	8.56	6.31	15.4
4	0.11	0.08	3.8

* indicates a rejection at the 5% level of significance

8 Conclusion

Preliminary analysis from fitting a multivariate structural time series model to the eleven Euro-zone countries indicates two possible convergence clubs, one a high income group, consisting of the five core economies plus Austria and Finland, and a low income group, made up of Spain, Greece and Portugal. Ireland seems to follow its own growth path. The multivariate convergence model is successful in separating trends from cycles and capturing the absolute convergence in the two groups. The evidence for convergence in the second group is less compelling, but the assumption of a single common trend is not unreasonable. The groups themselves appear to have converged in the relative sense. If this is correct, the implication is that the average per capita income in the poor group will remain almost 30% below that of the high group.

The cycles in the core high income group show a remarkable coherence in recent years, with the group standard deviation having fallen dramatically. There is less coherence in the poor group, though again there is evidence of a movement towards the same cycle as the rich group in recent years.

[14] This is appropriate (except possibly for $R = N - 1$) since the series are assumed to have drifts but no time trends in the co-integrating relationships. It could be argued that if clubs exhibit absolute convergence then constants should be excluded from the co-integrating relationships but this is not a standard constraint.

From the methodological point of view, the series illustrate the futility of trying to infer anything about convergence using Dickey-Fuller tests with constant included. However, it is possible to sensibly test against absolute convergence by dropping the constant and additional power is gained by pooling the information in several converging series. The likelihood ratio tests for co-integration appear to be of little use as a means of detecting the number of convergence clubs.

References

Barro, R. J., and X. Sala-i-Martin (1992): "Convergence," *Journal of Political Economy*, 100, 223-251.

Barro, R. J., and X. Sala-i-Martin (1995): *Economic Growth.* Boston, MA: McGraw-Hill.

Ben-David, D. (1993): "Equalizing exchange: trade liberalization and convergence," *Quarterly Journal of Economics*, 108, 653-79.

Bernard, A., and S. Durlauf (1995): "Convergence in international output," *Journal of Applied Econometrics*, 10, 97-108.

Bernard, A., and S. Durlauf (1996): "Interpreting tests of the convergence hypothesis," *Journal of Econometrics*, 71, 161-173.

Boldrin, M., and F. Canova (2001): "Inequality and convergence: reconsidering European regional economic policies," *Economic Policy*, 32, 207-53.

Carlino, G., and K. Sill (2001): "Regional income fluctuations: common trends and common cycles," *Review of Economics and Statistics*, 83.

Carvalho, V. M., and A. C. Harvey (2002): "Growth, cycles and convergence in US regional time series," DAE Working paper 0221, University of Cambridge.

Crafts, N., and G. Toniolo (1996): *Economic Growth in Europe since 1945.* Cambridge: Cambridge University Press.

Doornik, J. A. (1999): *Ox: An Object-Oriented Matrix Language.* 3rd edition. London: Timberlake Consultants Press.

Doornik, J. A., and D. F. Hendry (2000): *PcFIML.* London: Timberlake Consultants Ltd.

Durlauf, S. N., and P. A. Johnson (1995): "Multiple regimes and cross-country growth behaviour," *Journal of Applied Econometrics*, 10, 365-84.

Durlauf, S. N., and D. Quah (1999): "The new empirics of economic growth," in *Handbook of Macroeconomics, Volume 1, Chapter 4*, ed. by J. B. Taylor and M. Woodford. Elsevier Science.

Hamilton, J. D. (1994): *Time series analysis*, Princeton: Princeton University Press.

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

Harvey, A. C., and D. Bates (2002): "Multivariate unit root tests and testing for convergence," DAE Working paper 0301, University of Cambridge.

Harvey, A. C., and V. M. Carvalho (2002): "Models for converging economies," DAE Working paper 0216, University of Cambridge.

Harvey, A. C., and S. J. Koopman (1997): "Multivariate structural time series models," in *System dynamics in economic and financial models*, ed. by C. Heij et al.. Chichester: Wiley and Sons. pp. 269-298

Johansen, S. (1988): "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, 12, 13-54.

Koopman, S. J., A. C. Harvey, J. A. Doornik and N. Shephard (2000): *STAMP 6.0: Structural Time Series Analyser Modeller and Predictor.* London: Timberlake Consultants Ltd.

Koopman, S. J., N. Shephard and J. A. Doornik (1999): "Statistical algorithms for models in state space form using SsfPack 2.2," (with discussion) *Econometrics Journal*, 2, 107-160.

Martin, R. L. (2001): "EMU versus the Regions? Regional Convergence and Divergence in Euroland," *Journal of Economic Geography*, 1, 51-80.

Sala-i-Martin, X. (1996): "Regional cohesion: evidence and theories of regional growth and convergence," *European Economic Review*, 40, 1325-52.

Tsionas, M. G. (2000): "Real convergence in Europe: How robust are econometric inferences?," *Applied Economics*, 32, 1475-82.

Vahid, F., and R. F. Engle (1993): "Common trends and common cycles," *Journal of Applied Econometrics*, 8, 341-360.

EUROSTAT COLLOQUIUM
MODERN TOOLS FOR BUSINESS CYCLE ANALYSIS

Luxembourg, 28 and 29 November 2002
Session on "Cyclical Convergence and Forecasting"

FORECASTING MONTHLY MACROECONOMIC VARIABLES FOR THE EURO AREA

BY MASSIMILIANO MARCELLINO[1,2]

IEP- Bocconi Univerisity, IGIER and CEPR

We fit a variety of linear, non-linear and time-varying models to monthly Euro-area macroeconomic variables, obtained through two alternative aggregation schemes. Since non-linear models often over-fit in sample, we assess their relative performance in a real time forecasting framework. It turns out that in many cases linear models are beaten by non-linear specifications, a result that questions the use of standard linear methods for forecasting and modeling Euro-area variables.

KEYWORDS: European Monetary Union, Forecasting, Time-Varying models, Non-linear models, Instability, Non-linearity,

JEL CLASSIFICATION: C2, C53, E30.

[1]I am grateful to Jim Stock and Mark Watson for several conversations on the topics addressed in this paper, and for making available the GAUSS programs that form the basis of the code used in this paper. Participants at the Eurostat colloquium on "Modern tools for business cycle analysis" also provided valuable comments. The usual disclaimers apply.

[2]Correspondence to: IGIER, Università Bocconi, Via Salasco 3/5, 20136, Milan, Italy.
E-mail: massimiliano.marcellino@uni-bocconi.it Phone: +39-02-5836-3327 Fax: +39-02-5836-3302

1 Introduction

In general, complicated time series models tend to fit well in sample, better than linear specifications, because of their greater flexibility. Yet, often the resulting models are too specific for the particular estimation sample, and their good performance is not replicated out of sample, in a forecasting context. For example, Stock and Watson (1999) found only minor gains from complicated models for forecasting about 200 macroeconomic variables for the US, on average linear specifications were quite good, and pooled forecasts even better.

On the other hand, the fact that Euro-area series are an aggregate of national series which likely present structural breaks and nonlinear features because of the many changes that affected the European countries in the recent past indicates that nonlinear models could have a true advantage over simpler specifications.

The two previous considerations suggest to run the model comparison exercise, but ranking the models on the basis of their out of sample performance in a pseudo real time context. Marcellino (2002a, 2002b, 2002c) conducted a similar comparison but focusing either on single European countries or on quarterly Euro area series. The main result was that on average linear models or pooling methods were still the best, but non-linear specifications were forecasting well for a good percentage of the macroeconomic variables under analysis. Basically, while linear models forecast reasonably well for all the series, non-linear models are quite good for some series but quite bad for others, so that on average they are beaten by simpler methods.

In this paper we evaluate whether this is the case also for monthly Euro-area variables, obtained by two alternative aggregation procedures. Specifically, Marcellino, Stock and Watson (2001) use constant aggregation weights, proportional to the real GDP share in 1995, to construct monthly series for the period 1982:1-1997:8. The resulting series are rather close to the official Eurostat variables over the '90s. From their dataset, we analyze industrial production (IPMSW) and consumer prices (CPIMSW), while we do not consider the unemployment rate since the series starts later on. Instead, Beyer, Doornik and Hendry (2001) adopt a more sophisticated aggregation procedure, which is based on the aggregation of growth rates of the variables, with time-varying weights, whose evolution depends also on the behavior of the exchange rate. The Euro zone aggregate growth series are then cumulated to obtain the levels of the variables. The method yields real gdp (YERBDH), the gdp deflator (YEDBDH) and M3 (M3BDH), monthly, over the period 1980:1-1999:2.

We compare three main forecasting methods. The linear method, which includes autoregressive (AR) models, exponential smoothing and random walk models. The time-varing method, which includes time-varying AR models and smooth transition AR models. The non-linear method, that includes artificial neural network models. Within each method we consider several alternative specifications, for a total of 58 models, and different forecast horizons. We then add to the comparison three common forecast pooling procedures, that yields a total of 74 competing forecasts.

The forecasts are described in details in Section 2. Their relative performance is evaluated in Section 3 for the two datasets. Section 4 summarizes the main findings of the paper and offers suggestions for further extensions.

2 Forecasting methods

We consider forecasting models of the type

$$y_{t+h}^{h} = f\left(Z_t; \theta_{ht}\right) + \varepsilon_{t+h}, \tag{1}$$

where y_t is the variable being forecast, h indicates the forecast horizon, Z_t is a vector of predictor variables, ε_t is an error term, and θ_h is a vector of parameters, possibly evolving over time. Forecasting methods differ for the choice of the functional form of the relationship between y_{t+h} and Z_t, f. Within each method, different models are characterized by the choice of the regressors Z_t and the stationarity transformation applied to y_t.

The h-step forecast is

$$\widehat{y}_{t+h}^{h} = f\left(Z_t; \widehat{\theta}_{ht}\right), \tag{2}$$

with associated forecast error

$$e_{t+h} = y_{t+h}^{h} - \widehat{y}_{t+h}^{h}. \tag{3}$$

When y_t is treated as stationary, it is $y_{t+h}^{h} = y_{t+h}$, while if y_t is I(1) then $y_{t+h}^{h} = y_{t+h} - y_t$. Besides computing results for both cases, we also consider a pre-test forecast where the decision on the degree of integration of y_t depends on a unit root test. Pre-testing often improves the forecasting performance, see e.g. Diebold and Kilian (2000). Specifically, we use the Elliot, Rothenberg and Stock (1996) DF-GLS statistics, which performed best in the simulation experiments in Stock (1996). Note that $e_{t+h} = y_{t+h}^{h} - \widehat{y}_{t+h}^{h}$, independently of whether y_t is treated as stationary or not, so that forecast errors from the three different cases (stationary, I(1) and pre-test) can be directly compared.

To mimic real time situations, for each variable, method and model the pre-test for unit root, estimation, and model selection are repeated each month over the forecasting period.

Because of the short sample available, we consider forecast horizons, h, of 1, 3 and 6 months. When h is larger than one, the "h-step ahead projection" approach in (1), also called dynamic estimation (e.g. Clements and Hendry (1996)), differs from the standard approach of estimating a one-step ahead model, and iterate it forward to obtain h-step ahead predictions. The h-step ahead projection approach has two main advantages in this context. First, the impact of specification errors in the one-step ahead model can be reduced by using the same horizon for estimation as for forecasting. Second, simulation methods are not required to obtain forecasts from non-linear models. The resulting forecasts could be slightly less efficient, see e.g. Granger and Terasvirta (1993, Ch.8), but the computational savings in our real time exercise with many series and models are substantial.

In few cases there are problems with the estimation of the non-linear models, which then yield very large forecast errors. We introduced an automatic forecast trimming procedure, in order not to bias the comparison against these methods. In particular, when the absolute value of a forecasted change is larger than any previously observed change, a no change forecast is used.

Let us now describe first the forecasting methods and models, and then the pooling procedures we compare. More details can be found in Stock and Watson (1996, 1999), Marcellino (2002a, 2002b, 2002c).

Linear methods

Autoregression (AR). Though very simple, these models have performed rather well in forecast comparison exercises, see e.g. Meese and Geweke (1984), or Marcellino, Stock and Watson (2001) for the Euro area. The f function in (1) is linear, and Z_t includes lags of the y variable and a deterministic component. The latter can be either a constant or also a linear trend. The lag length is either fixed at 4, or it is chosen by AIC or BIC with a maximum of 6 lags. Given that the variable y_t can be treated as stationary, I(1), or pre-tested for unit roots, overall we have 18 models in this class.

Exponential smoothing (ES). Makridakis et al. (1982) found this method to perform rather well in practice even though, from a theoretical point of view, it is optimal in the mean square forecast error sense only when the underlying process follows a particular ARMA structure, see e.g. Granger and Newbold (1986, Ch.5). We consider both single and double exponential smoothing, which are usually adopted for, respectively, stationary and trending series. Estimation of the parameters is conducted by means of (recursive) non-linear least squares (see e.g. Tiao and Xu (1993)). The third model in this class is given by a combination of the single and double models, based on the outcome of the unit root test.

Non-linear methods

Time-varying autoregression (TVAR). Following Nyblom (1989), we let the parameters of the AR model evolve according to the following multivariate random walk model:

$$\theta_{ht} = \theta_{ht-1} + u_{ht}, \quad u_{ht} \sim iid\left(0, \lambda^2\sigma^2 Q\right) \tag{4}$$

where σ^2 is the variance of the error term ε_t in (1), $Q = \left(E\left(Z_t Z_t\right)\right)^{-1}$, and we inspect several values of λ: 0 (no variation), 0.0025, 0.005, 0.0075, 0.01, 0.015, or 0.020. We consider first a specification with a constant, 3 lags and λ= 0.005, and then we allow for AIC or BIC selection of the number of lags (1, 3 or 6) jointly with the value of λ. In each case, y_t can be either stationary, or I(1) or pre-tested, so that we have a total of 9 TVAR models. The models are estimated by the Kalman filter.

Logistic smooth transition autoregression (LSTAR). The generic LSTAR model can be written as

$$y_{t+h} = \alpha'\zeta_t + d_t\beta'\zeta_t + \varepsilon_{t+h}, \tag{5}$$

where $d_t = 1/\left(1 + \exp\left(\gamma_0 + \gamma_1\zeta_t\right)\right)$ and $\zeta_t = (1, y_t, y_{t-1}, \ldots, y_{t-p+1})$ if y_t is treated as stationary or $\zeta_t = (1, \Delta y_t, \Delta y_{t-1}, \ldots, \Delta y_{t-p+1})$ if y_t is I(1). The smoothing parameter γ_1 regulates the shape of parameter change over time. When $\gamma_1 = 0$ the model becomes linear, while for large values of γ_1 the model tends to a self-exciting threshold model, see e.g. Granger and Terasvirta (1993), Terasvirta (1998) for details. For models specified in levels we consider the following choices for the threshold variable in d_t: $\zeta_t = y_t$, $\zeta_t = y_{t-2}$, $\zeta_t = y_{t-5}$, $\zeta_t = y_t - y_{t-6}$, $\zeta_t = y_t - y_{t-12}$. For differenced variables, it can be $\zeta_t = \Delta y_t$, $\zeta_t = \Delta y_{t-2}$, $\zeta_t = \Delta y_{t-5}$, $\zeta_t = y_t - y_{t-6}$, $\zeta_t = y_t - y_{t-12}$. In each case the lag length of the model was either 1 or 3 or 6. We report results for the following models: 3 lags and $\zeta_t = y_t$ (or $\zeta_t = \Delta y_t$ for the I(1) case); 3 lags and $\zeta_t = y_{t-6}$; AIC or BIC selection of both the number of lags and the specification of ζ_t. In each case, y_t can be either stationary, or I(1) or pre-tested, so that overall there are 12 LSTAR models. Estimation is carried out by (recursive) non-linear least squares, using an optimizer developed by Stock and Watson (1999).

Artificial neural network (ANN). ANN models can provide a valid approximation for the generating mechanism of a vast class of non-linear processes, see e.g. Hornik, Stinchcombe and White (1989), and Swanson and White (1997) for their use as forecasting devices. The single layer feedforward neural network model with n1 hidden units (and a linear component) is specified as:

$$y_{t+h} = \beta_0' \zeta_t + \sum_{i=1}^{n_1} \gamma_{1i}' g\left(\beta_1' \zeta_t\right) + \varepsilon_{t+h}, \tag{6}$$

where $g(x)$ is the logistic function, $g(x) = 1/(1+e^x)$. Even more flexibility can be obtained with the double layer feedforward neural network with n_1 and n_2 hidden units:

$$y_{t+h} = \beta_0' \zeta_t + \sum_{j=1}^{n_2} \gamma_{2i} g\left(\sum_{i=1}^{n_1} \beta_{2ji} g\left(\beta_{1i}' \zeta_t\right)\right) + \varepsilon_{t+h}. \tag{7}$$

We report results for the following specifications: $n_1 = 2$, $n_2 = 0$, $p = 3$ (recall that p is number of lags in ζ_t); $n_1 = 2$, $n_2 = 1$, $p = 3$; $n_1 = 2$, $n_2 = 2$, $p = 3$; AIC or BIC selection with $n_1 = (1,2,3)$, $n_2 = (1,2$ with $n_1 = 2)$, $p = (1,3)$. For each case y_t can be either stationary, or I(1) or pre-tested, which yields a total of 15 ANN models. The models are estimated by (recursive) non-linear least squares, using an algorithm developed by Stock and Watson (1999).

No- change

No change (NC). The random walk based forecast is simply $\widehat{y}_{t+h} = y_t$. Notwithstanding its simplicity, in a few cases it was found to outperform even forecasts from large-scale structural models, see e.g. Artis and Marcellino (2001).

Pooling procedures

Linear combination forecasts (C). These forecasts are weighted averages of those described so far:

$$\widehat{y}_{t+h} = \sum_{m=1}^{M} k_{m,h,t} \widehat{y}_{t+h,m}, \qquad k_{m,h,t} = \left(1/msfe_{m,h,t}\right)^w / \sum_{j=1}^{M} \left(1/msfe_{j,h,t}\right)^w \tag{8}$$

where m indexes the models, $k_{m,h,t}$ denotes the weights, and $msfe$ indicates the mean square forecast error. Bates and Granger (1969) showed that the weighting scheme that minimizes the msfe of the pooled forecasts involves the covariance matrix of all the forecast errors, which is unfeasible in our case because M is very large. Hence, following their suggestion, the weight of a model is simply chosen as inversely proportional to its msfe, which is equivalent to setting $w = 1$ in equation (8). We also consider the cases $w = 0$, equal weight for each forecast, and $w = 5$, more weight for the best performing models. Moreover, we analyze separately pooling the linear models only, the non-linear models only, and all the models. Thus, overall we have 9 linear combination forecasts.

Median combination forecasts (M). These are the median forecasts from a set of models, and are computed because with non-Gaussian forecast errors linear combinations of the forecasts are no longer necessarily optimal. As in the previous method, we distinguish among three groups of models: linear, non-linear, and all models. Thus, we have 3 median combination forecasts.

Predictive least squares combination forecasts (PLS). In this approach the model is selected on the basis of its past forecasting performance over a certain period, that for us is one year. Thus, the model that produced the lowest msfe over the past year is used as the forecasting model, and the choice is recursively updated each month over the forecast period. We compute 4 of these forecasts, that differ for the set of models compared: all models, all linear models, all non-linear models, all models plus the linear and the median combination forecasts.

The 58 models and the 16 pooling procedures to be used in the forecast comparison exercise are summarized in Table 1.

3 Forecast comparison

To compare the competing forecasts we need to choose a loss function. We define the loss from forecast m for variable n as

$$Loss_{n,m}^{h} = \frac{1}{T-h} \sum_{t=1}^{T-h} |e_{t+h,n,m}|^{\rho}, \tag{9}$$

where e_{t+h} is the h-step ahead forecast error, and ρ can be equal to 1, 1.5, 2, 2.5 or 3. The values $\rho = 1$ and $\rho = 2$ correspond to the familiar choices of, respectively, the mean absolute error (mae) and the mean square forecast error (msfe) as the loss function. We can now use the loss function in (9) to rank the forecasts described in the previous section for the different variables of interest.

As mentioned in the Introduction, the BDH and MSW datasets are monthly. The former includes the GDP (YERBDH), the GDP deflator, and M3 (M3BDH), and the forecasting period is 1991:1-1999:2 with estimation starting in 1980. The latter includes the consumer price index (CPIMSW) and industrial production (IPMSW), and the forecasting period is 1993:1-1997:8 with estimation starting in 1982. In both cases the forecasting horizons we consider are $h = 1, 3, 6$ months.

Table 2 reports the best 5 forecasting models for the BDH variables. For YERBDH there is a clear preference for ANN models, with some pooled forecasts in the top-5 for longer forecast horizons. The performance of ANN models is also good for YEDBDH and M3BDH, though smooth transition models or pooling is preferred for YERBDH when $h = 1$ and for M3BDH when $h = 3, 6$. It is worth mentioning that in general the ranking of the models is not affected by the value of ρ.

The results for the MSW variables are summarized in Table 3. For CPI the outcome is similar to the BDH case, with pooling procedures ranked first for $h = 1$ and ANN models for $h = 3, 6$. For IP the results are more varied depending on the value of ρ, with the robust finding that linear models do not perform well.

The poor performance of the linear models is actually the most interesting finding of the analysis. The fact that nonlinear models are particularly good for the BDH dataset is likely due to the time-varying weights used in the aggregation process, but the good performance for the MSW series, which are obtained with constant weights, indicates that some of the national series also present nonlinear features, which is in line with the findings in Marcellino (2002a, 2002b).

4 Conclusions

In this paper we have provided a thorough analysis of the relative merits of linear, time-varying, non-linear and pooled forecasts for aggregate monthly Euro-area macroeconomic variables. The main result is that for several variables forecasts from linear models can be substantially improved upon.

This finding is interesting for policy purposes, but also more generally for empirical macroeconomic analysis. For example, it suggests that measures of persistence based on linear specifications can be inappropriate, as well as impulse response functions. Also, the use of GMM for the estimation of EMU forward looking Taylor rules is questionable, since the relationship between future values of inflation and the instruments can be non-linear or time-varying.

The main limitation of the current analysis is that it lacks a theoretical economic explanation for the results. The many changes that affected the economies now in the EMU can explain the failure of the linear model, but they do not provide a clear indication on the pattern of time variation of the parameter or the type of non-linearity to be included in the statistical models for the EMU variables. Further research in this area is required.

References

Artis, M., and M. Marcellino (2001): "Fiscal forecasting: the track record of IMF, OECD and EC," *Econometrics Journal*, 4, s20-s36.

Bates, J. M., and C. W. J. Granger (1969): "The combination of forecasts," *Operations Research Quarterly*, 20, 415-468.

Beyer, A., J. A. Doornik, and D. F. Hendry (2001): "Constructing historical Euro-zone data," *The Economic Journal*, 111, F102-21.

Clements, M. P., and D. F. Hendry (1996): "Multi-step estimation for forecasting," *Oxford Bulletin of Economics and Statistics*, 58, 657-684.

Clements, M. P., and D. F. Hendry (2001): "Pooling of forecasts," *Econometrics Journal*, (forthcoming).

Diebold, F. X., and L. Kilian (2000): "Unit Root Tests are Useful for Selecting Forecasting Models," *Journal of Business and Economic Statistics*, 18, 265-273.

Elliott, G., T. J. Rothenberg, and J. H. Stock(1996): "Efficient tests for an autoregressive unit root," *Econometrica*, 64, pp. 813-36.

Granger, C. W. J., and P. Newbold (1986): *Forecasting economic time series*, San Diego: Academic Press.

Granger, C. W. J., and T. Terasvirta(1993): *Modelling non-linear economic relationships*, Oxford: Oxford University Press.

Hornik, K., M. Stinchcombe, and H. White (1989): "Multilayer feedforward networks are universal approximators," *Neural Networks*, 2, pp. 359-66.

Marcellino, M. (2002a): "Instability and non-linearity in the EMU," CEPR WP 3312.

Marcellino, M. (2002b): "Forecast pooling for short time series of macroeconomic variables," CEPR WP 3313.

Marcellino, M. (2002c): "Forecasting EMU macroeconomic variables," *International Journal of Forecasting* (forthcoming).

Marcellino, M., J. H. Stock, and M. W. Watson(2001): "Macroeconomic forecasting in the Euro area: country specific versus Euro wide information," *European Economic Review* (forthcoming).

Makridakis, S., A. Anderson, R. Carbonne, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, R. Winkler (1982): "The accuracy of extrapolation (time series) methods: Results of a forecasting competition," *Journal of Forecasting*, 1, pp. 111-153.

Meese, R., and J. Geweke(1984): A comparison of autoregressive univariate forecasting procedures for macroeconomic time series," *Journal of Business and Economic Statistics*, 2, pp. 191-200.

Nyblom, J. (1989): "Testing for constancy of parameters over time," *Journal of the American Statistical Association*, 84, pp. 223-230.

Swanson, N. R., and H. White (1997): "A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks," *The Review of Economics and Statistics*, 79, pp. 540-550.

Terasvirta, T. (1998): "Modelling economic relationships with smooth transition regressions" in Ullah, A. and Giles, D.E.A. (eds.), Handbook of Applied Economic Statistics, New York: Marcel Dekker, pp. 507-552.

Tiao, G. C., and D. Xu (1993): "Robustness of maximum likelihood estimates for multi-step predictions: the exponential smoothing case," *Biometrika*, 80, pp. 623-641.

Stock, J. H. (1996): "VAR, error correction and pretest forecasts at long horizons," *Oxford Bulletin of Economics and Statistics*, 58, pp. 685-701.

Stock, J. H., and M. W. Watson(1996): "Evidence on structural instability in macroeconomic time series relations," *Journal of Business and Economic Statistics*, 14, pp. 11-30.

Stock, J. H., and M. W. Watson (1999): "A comparison of linear and non-linear univariate models for forecasting macroeconomic time series," in Engle, R. and White, R. (eds), Cointegration, causality, and forecasting: A festschrift in honor of Clive W.J. Granger, Oxford: Oxford University Press, pp. 1-44.

Table 1: Forecasting models

A. Linear methods

ARF(X,Y,Z) *Autoregressive models (18 models)*

X = C (const.) or T (trend)
Y = 0 (stationary), 1 (I(1)), P (pre-test)
Z = 4 (4 lags), a (AIC), b (BIC)

EX(X) *Exponential smoothing (3 models)*

X = 1 (single), 2 (double), P (pre-test)

B. Non-linear methods

ARTVF(X,Y,Z) *Time-varying AR models (9 models)*

X = C (const.)
Y = 0 (stationary), 1 (I(1)), P (pre-test)
Z = 3 (3 lags), a (AIC), b (BIC)

LS(X,Y,Z) *Logistic smooth transition (6 models)*

X = 0 (stationary), 1 (I(1)), P (pre-test)
Y = transition variable, 10 ( ), 06 ( )
Z = 3 (p, lag length)

LSF(X,W) *Logistic smooth transition (6 models)*

X = 0 (stationary), 1 (I(1)), P (pre-test)
W = a (AIC on transition variable and p), b (BIC)

AN(X,Y,Z,W) *Artificial neural network models (9 models)*

X = 0 (stationary), 1 (I(1)), P (pre-test)
Y = 2 (n1)
Z = 0, 1, 2 (n2)
W = 3 (p, lag length)

ANF(X,S) *Artificial neural network models (6 models)*

X = 0 (stationary), 1 (I(1)), P (pre-test)
S = a (AIC on n1, n2, p), b (BIC)

C. No Change

NOCHANGE *No change forecast (1 model)*

D. Pooling

C(X,Y) **Linear combination (9 forecasts)**

X = 1 (combine A,B,C), 2 (A only), 3 (B only)
Y = 0, 1, 5 (weight, w in equation (8))

M(X) *Median combination (3 forecasts)*

X = 1 (combine A,B,C), 2 (A only), 3 (B only)

P(X) *Predictive least square combination (4 forecasts)*

X = 1 (combine A,B,C), 2 (A only), 3 (B only), A (A,B,C,D)

Table 2: Best models for each variable with different loss functions, BDH dataset

Var/ρ		1.00	1.50	2.00	2.50	3.00
yerbdp	h=1	ANFPa	ANF1a	ANFPa	ANFPa	ANFPa
	h=3	ANF0b	ANF0b	ANF0b	ANF0b	ANF0b
	h=6	ANF0a	ANF0a	ANF0a	ANF0a	ANF0a
	h=1	ANF1a	ANFPa	ANF1a	ANF1a	ANF1a
	h=3	ANF0a	ANF0a	ANF0a	ANF0a	ANF0a
	h=6	ANF0b	ANF0b	ANF0b	ANF0b	ANF0b
	h=1	ANP223	ANP223	ANF1b	AN1223	LSF0b
	h=3	P3	P3	P1	P1	P1
	h=6	AN0213	AN0213	AN0213	AN0213	AN0213
	h=1	ANFPb	AN1223	AN1223	ANP223	AN1223
	h=3	P1	P1	P3	P3	P3
	h=6	P1	P1	P3	P3	P3
	h=1	AN1223	ANF1b	ANP223	ANF1b	ANF1b
	h=3	PA	PA	PA	PA	PA
	h=6	P3	P3	P1	P1	P1
yedbdh	h=1	LSF0a	LSF0a	LSF0a	LSF0a	LSF0a
	h=3	ANF0a	ANF0a	ANF0a	ANF0a	ANF0a
	h=6	ANF0a	ANF0a	ANF0a	ANF0a	ANF0a
	h=1	P3	P3	P3	P3	P3
	h=3	ANF0b	ANF0b	ANF0b	ANF0b	ANF0b
	h=6	ANF0b	ANF0b	ANF0b	ANF0b	ANF0b
	h=1	P1	P1	P1	P1	P1
	h=3	P1	P1	P1	P1	P1
	h=6	P1	P1	P1	P3	P3
	h=1	PA	PA	PA	PA	PA
	h=3	P3	P3	P3	P3	P3
	h=6	P3	P3	P3	AN0223	AN0223
	h=1	LSF0b	LSF0b	LSF0b	LSF0b	LSF0b
	h=3	PA	PA	PA	PA	PA
	h=6	AN0223	AN0223	AN0223	P1	P1
m3bdh	h=1	P1	AN0223	ANF0b	ANF0b	ANF0b
	h=3	LSF0b	LSF0b	LSF0b	LSF0b	LSF0b
	h=6	P1	ANF0a	P3	ANF0a	P1
	h=1	P3	P3	ANF0a	ANF0a	ANF0a
	h=3	LSF0a	LSF0a	LSF0a	LSF0a	LSF0a
	h=6	ANF0a	P3	P1	P1	P3
	h=1	AN0223	P1	P1	P3	P3
	h=3	ANF0a	ANF0a	C3001	C3001	C3001
	h=6	P3	P1	ANF0a	P3	ANF0a
	h=1	LSF0a	ANF0b	P3	P1	P1
	h=3	AN0213	AN0213	ANFPa	ANFPb	C1001
	h=6	AN0223	ANF0b	AN0223	AN0223	ANF0b
	h=1	LSF0b	ANF0a	AN0223	AN0223	AN0223
	h=3	ANF0b	ANF0b	ANF1a	ANF1b	ANF1a
	h=6	ANF0b	AN0223	ANF0b	ANF0b	AN0223

Notes: See Table 1 for the definition of the models. The loss function is $L^h_{n,m} = ((Loss^h_{n,m})/(Loss^h_{n,1})), Loss^h_{n,m} = (1/(T-h))\sum_{t=1}^{T-h}|e_{t+h,n,m}|^{\rho}$, where the benchmark model is ARFC04 and e_{t+h} is the h-step ahead forecast error.

Table 3: Best models for each variable with different loss functions, MSW dataset

Var/ρ		1.00	1.50	2.00	2.50	3.00
cpimsw	h=1	P3	PA	P3	AN0223	AN0223
	h=3	AN0203	ANF0a	ANF0a	ANF0a	ANF0a
	h=6	AN0223	AN0223	AN0223	AN0223	AN0223
	h=1	P1	P1	P1	P1	P3
	h=3	ANF0a	AN0223	AN0223	AN0223	AN0203
	h=6	ANF0b	ANF0b	ANF0b	ANF0b	ANF0b
	h=1	AN0223	AN0223	PA	P3	PA
	h=3	AN0223	AN0203	AN0203	AN0203	AN0223
	h=6	ANF0a	ANF0a	ANF0a	ANF0a	ANF0a
	h=1	PA	P3	AN0223	PA	P1
	h=3	ANF0b	ANF0b	ANF0b	ANF0b	ANF0b
	h=6	EX2	EX2	EX2	EX2	EXP
	h=1	AN0203	ANF0a	AN0203	ANF0a	AN0203
	h=3	EXP	EXP	EXP	EX2	EX2
	h=6	EXP	EXP	EXP	EXP	EX2
ipmsw	h=1	EXP	ANP223	LSFPa	LSF1a	LSP103
	h=3	C2000	C2000	C2000	ARTVFC1b	ARTVFCPb
	h=6	C1000	C1000	C2000	ARFT0b	ARFT04
	h=1	EX2	AN1223	LSF1a	LSP103	LS1103
	h=3	ARFCPa	C2001	C2001	ARTVFC1a	ARTVFC1b
	h=6	C1001	C1001	C2001	ARFT04	ARFT0b
	h=1	ARFT04	LSP103	LSP103	LSFPa	LSF1a
	h=3	ARFC14	ARFC14	ARTVFCPb	ARTVFCPb	ARTVFCPa
	h=6	C3001	C2000	LS1103	ARFT0a	ARFT0a
	h=1	AN1223	LS1103	LS1103	LS1103	LSFPa
	h=3	ARFCP4	ARFTPa	ARTVFC1a	ARTVFCPa	ARTVFC1a
	h=6	C3000	C2001	LSP103	C2000	LS1103
	h=1	ANP223	LSF1a	ANP223	AN1223	ANF1b
	h=3	ARFTP4	ARFCP4	ARTVFC1b	ARFTP4	ARFT04
	h=6	LS1063	C3001	C1001	LS1103	LSP103

Notes: See Table 1 for the definition of the models. The loss function is $L^h_{n,m} = ((Loss^h_{n,m})/(Loss^h_{n,1})), Loss^h_{n,m} = (1/(T-h))\sum_{t=1}^{T-h}(e_{t+h,n,m})^\rho$, where the benchmark model is ARFC04 and e_{t+h} is the h-step ahead forecast error.

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Cyclical Convergence and Forecasting"

IS THERE A COMMON EURO-ZONE BUSINESS CYCLE?

By James Mitchell and Kostas Mouratidis[1]

National Institute of Economic and Social Research, NIESR

This paper, using 40 years of monthly industrial production data, examines the relationship between the business cycles of the 12 Eurozone countries. We investigate both whether their business cycles have moved 'closer' together, and whether they have become more correlated over time. Since estimates of the business cycle have been found to be sensitive to how the cycle is measured, a range of alternative measures are considered. We focus on both parametric and nonparametric univariate measures of the 'classical' and 'growth' cycles.

[1]This paper is part of a programme of work undertaken at the National Institute of Economic and Social Research in London with the support of EUROSTAT. It develops the prior contributions of Massmann et al. (2002, 2003) and Massmann & Mitchell (2002*a*, 2002*b*). The views expressed in this paper, of course, are those of the authors alone. Address for correspondence: Dr. James Mitchell, NIESR, 2 Dean Trench Street, Smith Square, London, SW1P 3HE, UK. Tel: +44 (0)207 654 1926. E-Mail: j.mitchell@niesr.ac.uk.

1 Introduction

A popular interpretation of the traditional Burns-Mitchell definition of a business cycle is of comovement among individual economic variables. Although this interpretation has been challenged recently, see Harding and Pagan (2002b), the existence of a 'common' Euro-zone business cycle requires the individual business cycles of the 12 Euro-zone countries to be synchronised.

Optimists might like to believe there is an emerging common cycle, but empirical evidence for a common cycle certainly is mixed. Consider, for example, the controversy initiated by Artis and Zhang (1997, 1999) and Inklaar and de Haan (2001) - while Artis and Zhang (1997, 1999) conclude that Euro-zone business cycles have become more synchronised, Inklaar and de Haan (2001) using the same, but updated, data reach the opposite conclusion. The importance of establishing whether there is a common Euro-zone business cycle derives from the view that a common cycle is a prerequisite for the successful operation of economic and monetary union (EMU) in Europe, for example see Christodoulakis, Dimelis, and Kollintzas (1995). The design, and implementation, of monetary policy in the Euro-zone is complicated in the absence of a common Euro-zone business cycle.

In this paper we use over 40 years of monthly industrial production data to re-examine whether the business cycles of the 12 current Euro-zone economies have become more synchronised, or correlated, over time.[2] Following Massmann and Mitchell (2002b), we generalise previous work by (i) computing standard errors associated with the correlation coefficients by using a generalised method of moments estimator, (ii) considering a series of rolling windows, rather than simply computing bilateral correlation coefficients for a small number of sub-periods [Artis and Zhang (1999) consider two periods, Inklaar and de Haan (2001) four] and (iii) proposing a summary statistic to capture counties' bilateral correlation coefficients. Furthermore, and importantly given the controversy over the appropriate definition of a business cycle, see *inter alia* Harding and Pagan (2002b), we consider a range of alternative measures of the business cycle.

One prominent view is that an economic time series can be decomposed into the sum of trend and cyclical components, the cyclical component commonly referred to as the 'growth' cycle. However, disagreement remains over how the trend component should be identified and estimated; a range of parametric and non-parametric algorithms have been considered, see Canova (1998a, 1998b) and Massmann, Mitchell, and Weale (2002, 2003) for more details. Canova (1998a, 1998b) and Massmann and Mitchell (2002a, 2002b) find business cycle inference to be sensitive to the de-trending method employed. An alternative view recently re-expressed by Harding and Pagan (2002b) rejects the concept of a trend-cycle decomposition and, in line with the NBER and Burns-Mitchell, defines the 'classical' business cycle in terms of the turning points of the original data series. Crucially this method does not rely on the estimation and extraction of a trend series.

Since we can imagine the situation where Euro-zone business cycles are uncorrelated but are

[2] Alternative multivariate approaches, rather than examining the correlation between individual business cycles, test whether a common factor drives them, for further discussion see Vahid and Engle (1993), Forni, Hallin, Lippi, and Reichlin (2000) and Harding and Pagan (2000). The existence of a common factor implies individual business cycles are correlated.

moving 'closer' together, due to less pronounced cyclical volatility, we also examine whether Euro-zone business cycles have moved closer together over time. Specifically, we consider the root mean squared difference between the Euro-zone growth business cycles, expressed as a percentage of potential or trend output, for a series of rolling windows. This is important since even if Euro-zone business cycles are uncorrelated, and so on this basis there is no evidence for a common Euro-zone business cycle, we should expect the harmful effects of common policy to be mitigated if Euro-zone business cycles are converging in the sense of becoming closer. For a related exercise assessing the compatibility of the UK with Euro-zone business cycles, see Massmann and Mitchell (2002a).

The plan of the paper is as follows. Section 2 describes the data that we use in our empirical analysis. In Section 3 an overview is given of the alternative measures of both the 'growth' and 'classical' business cycle that we consider. Section 4 presents the metric we use to summarise whether Euro-zone business cycles are becoming increasingly correlated. Empirical results are presented in Section 5 and Section 6 concludes.

2 Data

The data used for our empirical analysis are the natural logarithms of the industrial production indices (IIP) of Austria, Belgium, Finland, France, Germany, Greece, Italy, Luxembourg, the Netherlands, Portugal and Spain, as compiled and seasonally adjusted by the OECD. The data set is an updated version of that employed by Artis and Zhang (1997, 1999)[3] as well as Inklaar and de Haan (2001). The data are monthly and the sample period is 1962-1 to 2002-5.

The use of industrial production data for business cycle analysis is justified by appealing to the historically strong correlation, importantly in growth rates given their trending nature, of industrial production and gross domestic product (GDP), the preferred measure of 'aggregate economic activity', see Harding and Pagan (2002b). In contrast to GDP data, monthly observations on industrial production are available on a consistent basis for most OECD countries back to the 1960s. With over 40 years of data we can meaningfully identify and estimate business cycles. Note, however, that there are reports that the historically close relationship between industrial production and GDP has weakened. If confirmed, this would clearly make it more debatable whether meaningful inference about GDP can be made.

Like Artis and Zhang, we also, in selected cases, consider a control group (the US) to ascertain whether any conclusion about the Euro-zone is spurious and is in fact a property of world business cycles and not specific to the Euro-zone.

We considered the impact outliers might have on the level of the data series. This is important because aberrant observations can distort inference. In particular, we used the automatic procedure of the TRAMO-SEATS software package, see Gomez and Maravall (1998), to detect and remove any outlying observations in the data series. Non-outlier corrected results, however, are qualitatively similar to those presented below.

[3] We are very grateful to Mike Artis for making the series available to us.

3 Identifying the business cycle

The conclusions of Artis and Zhang (1997, 1999) and Inklaar and de Haan (2001) about European business cycles make recourse only to a subset of the many measures of the business cycle. We consider a larger set of popular measures of the growth cycle, cf. Canova (1998a). Attention is restricted to univariate detrending methods. Additionally, we consider both parametric and nonparametric methods of identifying classical business cycles, based on an analysis of turning points. Let us consider the detrending algorithms, that yield the growth cycles, first.

3.1 'Growth' business cycles

Trend-cycle decompositions are usefully grouped into parametric and non-parametric de-trending algorithms.[4] The former are based on an estimated parametric statistical model, while the latter are motivated on a more *ad hoc* and intuitive basis.

3.1.1 Parametric measures of the growth cycle

Three parametric methods are considered: an ARIMA model resulting in the Beveridge-Nelson (BN) decomposition, an unobserved components model with a smooth stochastic trend, and a linear regression model with a constant time trend. The first two models are cast in state-space form and estimated using exact maximum likelihood. In particular, calculations make use of the SsfPack module for Ox, see Koopman, Shephard, and Doornik (1999) and Doornik (1998).

1. The Beveridge-Nelson (BN) decomposition. Beveridge and Nelson (1981) define the trend component as the limiting forecast of a stochastic process adjusted for the mean growth rate; the BN cycle is then the sum of the trend and the original series. There are a number of ways the BN cycle can be computed, see in particular the exact algorithms suggested by Newbold (1990), Arino and Newbold (1998) and Morley, Nelson, and Zivot (2002). We adopt the traditional algorithm, suggested by Beveridge and Nelson (1981), that involves an approximation that proved numerically almost identical to the exact result. Specifically we used the Akaike Information Criterion to choose the preferred multiplicative seasonal ARMA model of order $(p, q) \times (p_s, q_s)$, estimated in first differences of the level of the series, from the set $p, q = \{0, 1, 2, 3, 4\}$ and $p_s, q_s = \{0, 1, 2\}$. Across countries, we found the resulting BN cycle to be extremely volatile with low persistence, see Massmann and Mitchell (2002b) for details. The BN cycle did not accord at all with our prior of what a business cycle should look like, and as a result is not considered further.

2. An unobserved components (UC) model. Following the recent literature, see for instance Koopman, Shephard, and Doornik (1999, sc. 3.2), we estimate a smooth trend model that decomposes the process y_t into $y_t = \mu_t + c_t + \xi_t$ where μ_t, c_t, and ξ_t are the trend,

[4]This distinction is made purely for expositional reasons. It should be noted that non-parametric de-trending methods can be rationalised as parametric ones, see in particular the paragraphs on the Hodrick-Prescott and ideal band pass filters below. Furthermore, the parametric methods, like the non-parametric ones, are 'simply' taking weighted moving averages of the data, see Harvey and Koopman (2000).

cyclical and irregular components, respectively.[5] In particular, a smooth trend is obtained by setting $\sigma_\eta^2 = 0$ in the general local linear trend formulation:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad (1)$$
$$\beta_t = \beta_{t-1} + \zeta_t, \quad (2)$$

so that, effectively, $\eta_t = 0$ and the level of the trend, i.e. μ_t, is fixed while the trend's slope, β_t, remains stochastic since $\zeta_t \sim \mathsf{IID}\left(0, \sigma_\zeta^2\right)$. Moreover, it is assumed that the cyclical component c_t follows a trigonometric function, see for instance Harvey (1993, Section 6.5), and that the irregular component is $\xi_t \sim \mathsf{IID}\left(0, \sigma_\xi^2\right)$. The trend μ_t is then extracted from the data using the Kalman smoother. Importantly, the cycle of y_t is taken to be the observed series less the trend component, i.e. $y_t - \mu_t$, and thus consists of the sum of the irregular and the trigonometric cyclical components.

3. A linear regression (TIM) model. A univariate regression model is estimated, with an intercept and a constant linear trend as explanatory variables. The estimated cycle is given by the residuals.

3.1.2 Non-parametric measures of the growth cycle

Non-parametric methods offer an alternative means of extracting trend and cyclical components from a time series. In contrast to parametric methods they do not rely on estimation of a statistical model. For a survey, and examination of their theoretical properties within a unifying framework, see Massmann, Mitchell, and Weale (2002, 2003). We consider the following four non-parametric methods:

1. A centered moving average (MA).[6] Simple moving average detrending of time-series is used widely, see Osborn (1995). We consider a 4-year moving average.

2. The Hodrick-Prescott filter (HP).[7] Use of the Hodrick-Prescott filter requires one parameter, say λ, to be chosen, where λ controls the smoothness of the trend. Ravn and Uhlig (2002) argue that for monthly data we should set $\lambda = 129600$.[8]

3. An ideal band pass filter.[9] An ideal band pass filter is used to isolate the components of a time series that lie within a given range of frequencies. Economic theory can play a role in

[5] Morley, Nelson, and Zivot (2002) note that the BN cycle can be interpreted within the UC framework. UC models usually assume the disturbances to the trend and cyclical components are uncorrelated. When this restriction is relaxed the BN and UC approaches can yield identical cyclical estiimates.

[6] Other types of moving average filter are also used widely. For example, the Henderson moving average is used by the U.S. Bureau of Census' X-11 and X-12 procedures to extract the trend component of a time series, see Findley, Monsell, Bell, Otto, and Chen (1998).

[7] It is of theoretical interest that the Hodrick-Prescott filter also can be rationalised as a parametric method, since it has been interpreted by Harvey and Jaeger (1993) as the optimal estimator in an unobserved components time series model.

[8] When y_t is stationary Pedersen (2001) derives the optimal estimator of λ. Most economic time-series, however, contain a unit root.

[9] Again it is of theoretical interest that ideal band pass filters can be rationalised as optimal estimators in unobserved components time series models, see Harvey and Trimbur (2002).

defining these frequencies. In particular, given our interest lies in extracting the periodic components of an economic time series that can be associated with the business cycle, the bands can be chosen consistent with priors about the duration of the business cycle. For example, it is widely believed that a business cycle lasts between 1.5 and 8 years; the lower band can then be set at 18 months and the upper band at 96 months.[10] This removes low frequency trend variation and smooths high frequency irregular variation, while retaining the major features of business cycles. Since the ideal band pass filter requires a moving average of infinite order, in practice an approximation is required. We consider two approximations, firstly the widely used Baxter and King (1999) (BK) filter, and secondly the Christiano and Fitzgerald (1999) (CF) filter.[11] For the Baxter and King approximation we follow Baxter and King (1999) and set the length of the moving average at 3 years. For the Christiano-Fitzgerald approximation we follow their recommendation and drop two years data from the beginning and end of the filtered series.

4. The Phase Average Trend (PAT). The PAT involves a number of steps. First, compute deviations of the series from a centered moving average. Second, break up the deviations into phases, according to the dates of cyclical peaks and troughs. Third, compute the mean values of the series for each successive phase, and smooth using three-item or two-item moving averages. The PAT is then given by connecting these midpoints of these triplets or doublets, see Zarnowitz and Ozyildirim (2002). Following Artis and Zhang (1999) we use the PAT calculated by the OECD.[12]

3.2 'Classical' business cycles

Cycles computed in the manner discussed are classified as "growth cycles", since they are derived by deducting a trend, however estimated, from the original series. In contrast, "classical cycles" are defined in terms of the turning points in the levels of the original series. For example, a peak turning point indicates the end of an expansionary phase and the beginning of a recessionary one. This approach is typically associated with the National Bureau of Economic Research (NBER) and Burns and Mitchell (1946). There remains considerable controversy, and confusion, however, over what the "real" definition of a business cycle is, some arguing that it is the growth cycle, others that it is the classical cycle. For a detailed account of the arguments, see, for instance, Backhouse and Salanti (2000, pp.79-81).

Both parametric and nonparametric methods have been used to identify turning points. For further comparison of these two methods, see Harding and Pagan (2002a).

[10] Agresti and Mojon (2001) argue that 8 years is too low for the upper band; they argue that business cycles in Europe tend to last longer. In fact our results do not support such an argument, see Table 1-12 below. In any case, results using the Baxter-King filter were robust to increasing the upper band to 10 years.

[11] The approximation proposed by Baxter and King (1999) has been shown to fail to filter out the desired components when y_t is nonstationary, see Murray (2001) and Benati (2001). Christiano and Fitzgerald (1999) note that the optimal approximation to the ideal band pass filter requires knowing the true DGP of y_t. They derive the optimal approximation when y_t is $I(1)$. They find this particular approximation works well for standard macroeconomic time series.

[12] Estimates for the PAT are continually updated by the OECD as new data become available. Preliminary analysis using the most recent estimates, kindly made available to us by Ronny Nilsson of the OECD subsequent to the writing of this paper, suggests that the PAT results are qualitatively different from those presented below and, in fact, are closer to the results presented here for HP than TIM.

3.2.1 Parametric measures of the classical cycle

Following Hamilton's (1989) finding that Markov switching (MS) models can be used to date accurately the US business cycle, relative to NBER outcomes, MS models have been used widely to identify the classical business cycle.[13] MS models allow the growth rate of output to be influenced by an unobserved random variable that defines what state or regime the process is in at time t.

The original Hamilton MS model for the growth rate of output Δy_t is a fourth order autoregression with shifts in the mean:

$$(\Delta y_t - \mu_{st}) = \rho_1(\Delta y - \mu_{st-1}) + ... + \rho_4(\Delta y - \mu_{st-1}) + u_t, \tag{3}$$

where $u_t \sim i.i.d.N(0, \sigma^2)$ and the conditional mean μ_{st} is state dependent such that

$$\mu_{st} = \left\{ \begin{array}{l} \mu_0 \text{ if } s_t = 0 \text{ (i.e. recession or contraction)} \\ \mu_1 \text{ if } s_t = 1 \text{ (i.e. expansion)} \end{array} \right\} \tag{4}$$

and s_t, the unobserved state variable, follows a 2, or in general N, state ergodic Markov chain with transition probabilities $p_{ij} = \Pr(s_t = j|s_{t-1} = i)$, $\sum_{j=0}^{N-1} p_{ij} = 1$, $\forall i, j \in \{0, ..., N-1\}$. Estimation facilitates the computation of either the filtered, $\Pr(s_t = 1|\Delta y_1, ..., \Delta y_t)$, or smoothed, $\Pr(s_t = 1|\Delta y_1, ..., \Delta y_T)$, probabilities of an expansion, and these values, given a threshold value (typically 0.5), can be used to divide the sample into expansion and contraction periods.[14] Thus the data can be segmented into a binary series, b_t, that takes the value unity in expansions and zero in contractions. Summary statistics about the business cycle then can be compiled contingent on identification of the turning points, see Harding and Pagan (2001, 2002a). Interestingly, based on the segmented states b_t, the mean duration of an expansion (TP: trough to peak) is

$$E_b = \frac{\sum_{t=1}^{T} b_t}{\sum_{t=1}^{T-1}(1 - b_{t+1})b_t}, \tag{5}$$

and the mean duration of a contraction (PT: peak to trough) is

$$C_b = \frac{\sum_{t=1}^{T}(1 - b_t)}{\sum_{t=1}^{T-1}(1 - b_{t-1})b_t}. \tag{6}$$

By contrast, an alternative measure of the duration of an expansion (or contraction), based on the unobserved states, is given by

$$E_u = 1/(1 - p_{11}), \ (or \ C_u = 1/(1 - p_{00})). \tag{7}$$

Inference based on the segmented and unobserved states need not be the same.

In the application we consider two MS models, a constant transition probability model and a time-varying transition probability model. Preliminary analysis for the 12 Euro-zone countries

[13]It is of interest that under appropriate identification restrictions MS models also can be used to estimate the 'growth' business cycle. Lam's (1990) and Kim's (1994) extended version of Hamilton's (1989) model allows the explicit identification and estimation of the trend and cyclical components of the underlying series. For a review see Massmann, Mitchell, and Weale (2002, 2003).

[14]In our application the smoothed and filtered probabilities are almost identical and we report only the filtered probabilities.

confirmed the need to smooth the data, using a seven-month moving average, prior to analysis.[15] Smoothing both increased the duration of the business cycle from implausibly low values, and led to clearer evidence for switches in the mean, see (3). For similar reasons estimation was over the reduced sample period 1970m1-2002m5.

Constant transition probability model Two-state MS models, see (3), are considered, where both the intercept and variance are allowed to switch.[16] Allowing the variance to switch lets the volatility of the series vary according to the state or regime. The preferred MS model took the form:

$$\Delta y_t = \mu_{st} + \sigma^2_{st},\ s_t = 0 \text{ denotes contraction and } s_t = 1 \text{ denotes expansion} \quad (8)$$

Experimentation with alternative MS models, such as those with autoregressive lags (e.g..$\Delta y_t - \mu_{st} = \sum_{i=1}^{2}(\Delta y_{t-1} - \mu_{st-1}) + \sigma^2_{st}$), led to high transition probabilities, implying business cycles of implausibly short duration. Furthermore the filtered probabilities were highly erratic and inconsistent with empirical evidence for the growth rates.[17] This lends support to the specification (8) that has no autoregressive terms. However, as argued by Garcia and Perron (1996) spurious autoregressive terms can increase the value of the transition probabilities, but they suggest that this might be due to the presence of a third state. As discussed in footnote 15 preliminary results indicate that two states are appropriate in our case.

Time-varying transition probability model Although models of the form (3) and (8) allow the expected duration of expansions and contractions to differ, their durations are assumed to remain constant across time. Therefore, these models cannot capture the fact that the probability of moving from a recession into an expansion may depend on how long the economy has been in recession. Consequently, MS models with duration dependence have been proposed, e.g. see the univariate model of Durland and McCurdy (1994) that lets the probability of a regime shift be a function of the previous state, s_{t-1}, as well as its duration. Multivariate extensions by Filardo (1994, 1998) let the probability of a regime shift be a function of macroeconomic variables. However, an endogeneity problem might arise between the macroeconomic variables and the unobserved state variable.[18] One possible solution is suggested by Chourdakis and Tzavalis

[15] Previous work using monthly IIP data has also smoothed the data in an equivalent manner, see Artis, Krolzig, and Toro (1999).

[16] We considered a formal statistical test for the number of states, specifically Hansen's (1992) and Hansen's (1996) standardised likelihood ratio test. For simple MS models of the form (8) we tested the null hypothesis of linearity against the alternative of a two state MS model. For all countries the test provided strong evidence against linearity. Future work will extend the tests to consider three state MS models. Use will be made of a computationally less demanding test based on the ARMA representation of MS models, see Psaradakis and Spagnolo (2003). Artis, Krolzig, and Toro (1999) argue, also using monthly IIP data albeit over a different sample period, that three-states are appropriate. However, preliminary analysis using our data indicated that there was not a clear statistical difference between the 'low' and 'high' (positive) growth states indicating that two states may be appropriate.

[17] We considered also models where the autoregressive parameters are regime dependent. However, the transition probabilities were, in general, statistically insignificant.

[18] Filardo (1998) shows that when modelling time-varying transition probabilities the macroeconomic variables that affect time variation must be conditionally uncorrelated with the unobserved state variable. He suggests a solution based on joint estimation of the macroeconomic variables and the dependent variable.

(2000). Under Chourdakis and Tzavalis's (2000) specification the transition probabilities are determined as follows:

$$p_{ij} = \exp(\gamma_{0.ij} + \varphi_{ij} p_{ij}\{1\} + \gamma_{ij} u_{t-1}/[1 + \exp(\gamma_{0.ij} + \varphi_{ij} p_{ij}\{1\} + \gamma_{ij} u_{t-1})] \text{ for } i, j \in \{0, 1\} \quad (9)$$

where $p_{ij}\{1\}$ is the lagged value of the transition probability. $p_{ij}\{1\}$ is the duration dependence effect; it lets the probability of switching to another regime potentially increase the longer the economy has been in the current regime. The presence of u_{t-1} can represent the effects of exogenous, say policy, shocks on the transition probabilities. We allow for duration dependence following Chourdakis and Tzavalis (2000).

3.2.2 Nonparametric measures of the classical cycle

Following the NBER tradition, Harding and Pagan (2002b) advocate the use of a nonparametric dating rule to isolate turning points in the series. They suggest that the following three criteria need to be satisfied by the algorithm: (i) it picks the peaks and the troughs of a series, (ii) it ensures that peaks and troughs alternate and (iii) the cycle it defines has a minimum duration.

While this approach of characterising cycles has the attractive feature of not being dependent on removing trends estimated by some detrending algorithm, it does also require some parameters to be chosen. The censoring rule that specifies the minimum duration of the cycle must be stipulated. We follow Harding and Pagan and specify the censoring rule such that phases last at least two quarters (six months) and completed cycles (peak to peak, or trough to trough) last at least five quarters (15 months). Harding and Pagan show that if the dating algorithm is specified in this fashion it does a good job at approximating the turning points in quarterly U.S. GDP, as identified by the NBER, cf. Bry and Boschan (1971).

Therefore, in our setting, the process $\{y_t\}$ has a peak at time t if both of the following criteria are met:

$$\left.\begin{array}{l} \Delta_i y_t > 0, \\ \Delta_i y_{t+i} < 0, \end{array}\right\} \text{ for all } i = 1, \ldots, 6, \quad (10)$$

i.e. if at time t the series has a local maximum relative to the six months on either side. A trough is defined conversely. Application of the Harding-Pagan rule to the level of the series, $\{y_t\}$, yields a binary series with unity indicating a state of expansion, and zero a state of contraction.[19]

3.2.3 Comparing 'growth' and 'classical' business cycles

One implication of having identified the turning points of $\{y_t\}$ is that the information can be used to rank alternative methods of extracting a growth cycle according to their ability to match the turning points of the underlying time series, $\{y_t\}$. This provides an empirically based, sample specific, means of evaluating one detrending algorithm over another. Specifically, the turning points of $\{y_t\}$ can be treated as a benchmark, similarly to NBER turning points

[19]We would like to thank Don Harding for sending us GAUSS files to implement the Harding-Pagan rule.

in the U.S.. Then, like Canova (1999) for NBER turning points, we can see how well the alternative detrending methods do at matching this benchmark.[20]

The turning points of the alternative cyclical components are identified following Harding and Pagan. Their ability to match the turning points of $\{y_t\}$ is summarised by the correlation coefficient (C) between the two binary series. Robust standard errors for C, $se(C)$, are estimated by the generalised method of moments (GMM).

4 Correlation between Euro-zone business cycles

Following Massmann and Mitchell (2002b) consider n (in our case we have $n = 12$) countries' growth/classical business cycles identified *via* either a detrending method or a turning point rule.[21] At time t, $t = 1, ..., T$, estimate using GMM the N vector of contemporaneous correlation coefficients, ρ_t, $\{\rho_{jt}\}$, $j = 1, ..., N = n(n-1)/2$, where ρ_t is defined by stacking the columns of the lower triangular component of the $n \times n$ contemporaneous correlation matrix between the n cycles for a time-span of length h. Let $\widehat{\rho}_t$ denote the GMM estimate, $V(\widehat{\rho}_{j,t})$ its estimated variance and $cov(\widehat{\rho}_{jt}, \widehat{\rho}_{j',t})$ the covariance between $\widehat{\rho}_{j,t}$ and $\widehat{\rho}_{j',t}$, $j \neq j'$. The Newey and West (1987) nonparametric estimator, where the truncation parameter is set equal to $4(h/100)^{2/9}$, is used to estimate consistently the asymptotic variance matrix for the mean of the N moment conditions, even in the presence of serial correlation and heteroscedasticty. It should be noted that one could consider also leading and lagged correlation to examine whether certain countries' business cycles are leading or lagging others.

To summarise the N bilateral contemporaneous correlation coefficients we consider their mean, m_t, where

$$m_t = N^{-1} \sum_{j=1}^{N} \widehat{\rho}_{j,t}. \tag{11}$$

In fact, we consider a weighted mean, where the bilateral correlation coefficients are weighted according to the combined 'size' of the two countries under consideration. Size is proxied by total population in 1985. Weights derived from PPP GDP at constant prices proved qualitatively similar.

The estimated variance of the sample mean, $V(m_t)$, used to obtain confidence bands around m_t, is given by

$$V(m_t) = N^{-2} \left[\sum_{j=1}^{N} V(\widehat{\rho}_{j,t}) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} 2cov(\widehat{\rho}_{i,t}, \widehat{\rho}_{j,t}) \right] \tag{12}$$

A rise in m_t across time is viewed as evidence for increased synchronisation of business cycles consistent with the emergence of a 'common' Euro-zone business cycle. It should be noted that this is a necessary but not sufficient condition for a 'common' Euro-zone business cycle.

[20] Although widely used this benchmark may be considered theoretically flawed since the benchmark to which the growth cycle is compared is the classical cycle. There may be no theoretical reason to expect the two types of cycles to exhibit similar turning point behaviour.

[21] Using a turning point rule Harding and Pagan (2001), in the two-country case, propose a related measure of *synchronisation*. Whereas their statistical tests are based on *robust* regression coefficients, we consider *robust* correlation coefficients.

Massmann and Mitchell (2002b) note that it is important also to consider the variance of the correlation coefficients, as well as measures of the extent of intra-distributional movement within the distribution of correlation coefficients across time.

5 Empirical results

5.1 Alternative estimates of Euro-zone 'growth' and 'classical' business cycles

We have seen that there are many ways to measure a business cycle, but does this actually affect the way Euro-zone business cycles look? Let us first present some summary statistics that describe the properties of the alternative measures of Euro-zone business cycles. Then we will examine further the differences across these alternative measures by computing the full-sample correlation matrix across the 12 Euro-zone business cycles for each measure. This lets us ascertain to what extent the growth cycles estimated by the seven alternative de-trending methods, and the classical cycles derived *via* MS models and the Harding-Pagan rule, have fundamentally different statistical properties. As has been already noted, Canova (1998a, 1998b) found important differences across 'growth' cycles.

For each of the seven alternative estimates of the growth cycle, Tables 1-12 present two sets of summary statistics. The first set detail the statistical properties of the generated cycles, while the second set list their turning point characteristics identified by applying the Harding-Pagan rule to them. Note that data on the PAT are available up to 2001m8 only, and are not available for Belgium prior to 1970 and as a result we do not report PAT cycles for Belgium.

Consider the statistical properties of the growth cycles first. It is evident that there are important differences across de-trending methods. Take, for instance, the measure of dispersion, i.e. the standard deviation. The TIM and, to a lesser extent, PAT cycles display most variation, with standard deviations greater than 0.5, and are relatively persistent. By contrast the other cycles display less variation and provide more noisy estimates of the cycle. The period of the cycle, derived from the spectrum, varies too. The period is longer the more volatile the cycle. Indeed, the cycle with the highest variation according to this measure of the duration of a cycle, i.e. TIM, does not even cycle in the sense that its spectral density peaks at frequency zero.

Turning to the set of statistics based on application of the Harding-Pagan rule, we note that the growth cycles also differ with respect to the number of turning points detected. For instance, the cycles generated by the MA filter display far fewer turning points than the other de-trending methods even those, such as the BK and CF cycles computed over comparable sample-periods. However, a property common across many de-trending methods is that business cycles are asymmetric, in the sense that expansions last longer than contractions. The duration from peak to trough, and *vice-versa*, is an alternative measure of the cycle's period. There is no theoretical relationship, however, between this measure and that based on the cycles' spectra mentioned in the previous paragraph.

The final two columns of Tables 1-12 present the correlation coefficient C, and its estimated standard error, between the turning points of the seven estimates of the growth cycle and the classical cycle identified using the Harding-Pagan rule. They indicate that the turning points of the growth cycles, in general, explain a statistically significant amount of the variation in the classical cycle.

Table 13 presents the mean duration in months of business cycle phases for the three measures of the classical business cycle considered. Since the MS models are estimated over a shorter sample period (1970m1-2002m5) than considered for the 'growth' cycles, statistics for the Harding-Pagan measure also are presented for this period. For the MS model with constant transition probabilities measures of duration based on both the segmented and unobserved states are presented, see (5)-(7).

Using the Harding-Pagan rule the mean duration of a Euro-zone business cycle is seen in Table 13 to be approximately 60 months or 5 years, and expansions last much longer than contractions. This is broadly comparable to the measures of duration for the growth business cycles seen in Tables 1-12, although substantial differences again lie in the country-specific details. However, the MS models suggest quite different durations for the business cycle.

The business cycles implied by the MS models, in general, are shorter and more volatile than those for either the Harding-Pagan measure or the growth cycles. This seems to be a property shared by other MS models estimated for monthly IIP, see Artis, Krolzig, and Toro (1999). Moreover, the estimates of business cycle duration are sensitive to both the type of MS model considered as well as the measure of duration considered for the MS model.

Estimation results for the MS models did indicate that the transition probabilities are not constant over time. Plotting the time-varying transition probabilities revealed considerable variation over time. Furthermore, estimates of γ_{ij} and ϕ_{ij} were statistically significant in most cases again implying that transition probabilities are not constant across time; estimates suggested that the longer one stays in one regime (expansion or contraction) the more likely it becomes that one will switch regime. It should be noted that for a minority of countries, such as Ireland and Portugal, estimates indicated (as $\widehat{\mu}_0$ and $\widehat{\mu}_1$ were positive) that the states should not be interpreted as expansion and contraction, but rather 'high' and 'low' (positive) growth. For these countries we should re-interpret the PT phase as 'low' growth. The finding using the MS models, for say Portugal, that PT is greater than TP therefore is not saying recessions last longer than expansions. Results also indicate, in general, that the variance in a contraction is higher than that in an expansion. Further estimation results are available from the authors upon request.

Tables 14-23 present the matrices of correlation coefficients between the 12 Euro-zone business cycles for each measure of the business cycle. Estimated standard errors are in parentheses. The standard errors are estimated using GMM, and are consistent even in the presence of serial correlation and heteroscedasticity. Summarising the information in the Tables, correlation between countries varies dramatically according to the measure of the business cycle used. In general, correlation is highest for those measures of the 'growth' business cycle that were found in Tables 1-12 to have the highest standard deviation (low volatility), such as TIM. As a specific example consider the correlation between France and Germany, two of the Euro-zone's largest economies. Correlation ranges from 0.09 (for the MS model with constant transition probabilities) to 0.912 (for TIM), with the other measures in the range 0.33 (for the MS model with time-varying transition probabilities) to 0.67 (for MA). Since even a high correlation coefficient of, say, 0.9 between two cycles implies that one cycle explains just over 80% of the variation in the other cycle, nearly 20% of this variation is still left unexplained. These differences between alternative measures of the Euro-zone's business cycle could prove important in ascertaining whether there is a 'common' cycle, and it is clearly advisable not to restrict attention to just one measure of the business cycle unless one has a strong preference for one measure over another.[22]

5.2 Synchronisation of Euro-zone business cycles

Having established that it is important how we measure the business cycle, Figures 1-4 display for each measure the size-weighted average contemporaneous correlation, m_t cf. (11), between the 12 Eurozone business cycles for a 3-year rolling window centered on the mid-point of the window.[23] The 95% confidence bands, the dashed outer lines, provide a measure of uncertainty associated with the estimates.[24] For the MS models we consider the contemporaneous correlation between the filtered probabilities of an expansion, while for the Harding-Pagan measure we consider the correlation between the binary series, with unity denoting expansion.

Figures 1-4 reveal that correlation on average is positive, and in a statistically significant manner. Re-assuringly, Euro-zone business cycles are positively correlated. But there has been considerable volatility. Nevertheless, we can extract common features across the alternative measures of the business cycle. These include correlation trending upwards until the mid 1970s and reaching peaks of 0.4 to 0.75, according to the measure of the business cycle. Then, with for some measures a short-lived rise in the early 1980s, correlation in general falls to zero in the mid 1980s and is statistically insignificant. It is quite striking that this result is common across the alternative measures. Correlation then rises post mid 1980s to values in the range 0.25 to 0.5, before slumping quite rapidly around 1990. Since 1990 correlation between the Euro-zone countries has risen but remained volatile, experiencing a trough in the mid 1990s. Correlation has risen in the late 1990s to values in the range 0.25 to 0.5, and the most recent estimates indicate that it continues to rise. Without forecasting, it is, of course, too early to tell whether these recent trends will continue.

Although in this paper we seek to establish the *facts* rather than explain them, it is interesting to relate this behaviour to the exchange-rate regime. An interesting hypothesis is whether rises in the mean correlation coefficient coincide with periods when the exchange-rate has been fixed, to some degree.[25] This hypothesis has attracted considerable attention, see Artis and Zhang (1997, 1999) and Inklaar and de Haan (2001). Consistent with the findings of Artis and Zhang (1999) there is some supportive evidence in Figures 1-4. For example, the fall in correlation that occurred in the mid 1970s was in the aftermath of the collapse of the Bretton Woods fixed exchange-rate regime. Then the rise in correlation from the mid 1980s, and again from

[22] Arguably given that unit root tests indicate that the IIP data contain a unit root, the TIM cycles are our least preferred measure of the business cycle. A simple deterministic time trend cannot account for this nonstationarity even if it is due to a deterministic (structural) shift, rather than a unit root.

[23] Due to the possibility of long sequences of zeros or ones, it proved possible only to consider a longer window of 168 months (14 years) for the Harding-Pagan measure. Otherwise there may be no cyclical variation.

[24] Estimates of $V(m_t)$ are computed assuming independence, i.e. the covariances in (12) are set equal to zero. This avoided singularities in the variance-covariance matrix.

[25] Many other explanations for increased comovement between business cycles have been given. Increased correlation, for example, also has been related to international trade and geographical proximity, see Frankel and Rose (1998) and Clark and van Wincoop (2001), respectively.

the mid 1990s, was at a time when the EMS was relatively stable and credible; there were, for example, no exchange-rate re-alignments either in the late 1980s or from the mid 1990s, consistent with the entry requirements for EMU. By contrast the period of falling correlation in the early 1980s (specifically 1981-1986) was characterised by eleven re-alignments, see Gros and Thygesen (1998) for a more detailed chronology. The collapse in average correlation in the early 1990s can be explained away by German unification and the ensuing currency crises in 1992. These events, again see Gros and Thygesen (1998) for more details, temporarily disrupted the emergence of a 'common' Euro-zone business cycle. Whether or not one subscribes to this story, and more work certainly is needed, it is encouraging that in the run up to monetary union, and irrevocably fixed exchange-rates, correlation has been rising.

Finally we examine whether the recent evidence for more synchronised Euro-zone business cycles is specific to the Euro-zone or a tendency of world business cycles. As the control, we consider the Euro-zone's size-weighted average correlation with the US. We exclude Germany from the Euro-zone, as we wish also to consider these 11 economies' size-weighted average correlation with Germany. This is instructive, as increased correlation with Germany often is taken as evidence for a common Euro-zone business cycle. To summarise the bilateral correlations of Germany and the US against the 11 Euro-zone business cycles, we consider the size-weighted average contemporaneous correlation coefficient between them. This statistic is again computed for a series of rolling windows with a length of 3 years, except for the Harding-Pagan measure where a window length of 14 years is considered. Figures 5 and 6 present the results; for expositional ease 95% confidence bands are not presented.

Figure 5 presents the average correlation between Germany and the remaining 11 Euro-zone countries for each measure of the business cycle. Results, in general, are consistent with those in Figures 1-4. The most recent estimates again suggest that Euro-zone business cycles are becoming increasingly in sink with those of Germany.

Figure 6 then considers the average correlation of the Euro-zone (excluding Germany for consistency) with the US. Results indicate that this Euro-zone group is better correlated with Germany than the US. For example, the mean of the average correlation figures (presented in Figures 5 and 6) for each measure of the business cycle is higher over the period 1995m1-2000m2, and 1998m1-2000m2, against Germany than the US. Table 24 summarises the findings. Although correlation is higher against the US for the period 1998m1-2000m2 than 1995m1-2000m2, correlation also has risen against Germany. Moreover, correlation always is higher against Germany than the US. This supports the view that there is an emerging 'common' Euro-zone business cycle.

5.3 'Closeness' between Euro-zone business cycles

Euro-zone business cycles can be *compatible*, or be seen to be *converging*, even if they are not becoming increasingly correlated. This can occur if Euro-zone business cycles move closer together, in the sense that the cyclical disparity between them declines. Any reduction need not be associated with increased correlation.

One possible measure of closeness is to consider the root mean squared difference between the Euro-zone 'growth' business cycles, expressed as a percentage of potential or trend output, for a series of rolling windows. Since quantitative estimates of the 'output gap' are required we cannot conduct this exercise for the classical measures of the business cycle.

Figure 7 plots the size-weighted average root mean squared difference between Euro-zone business cycles for the alternative measures of the growth business cycle, for windows of 3 years. All measures bar two indicate that Euro-zone business cycles have moved closer together over time; cyclical disparity between the Euro-zone economies has declined to historically low levels.

We can see a slight increase in disparity in the mid 1980s and in the aftermath of German unification at the time of the currency crises in 1992, but the general trend clearly is downward. The TIM and PAT measures, however, indicate that despite declining from high levels in the 1970s, disparity within the Euro-zone has risen in the late 1990s. The other measures do not indicate such a pronounced rise over this late period. We believe the PAT and TIM measures are overstating the extent of the rise, since the trend component of both measures often fails to keep up with the increase in output. This leads to inflated estimates of the 'output gap'. By necessity our statistical approach to identifying and estimating business cycle relationships often requires a number of lead observations. This can prevent us examining the current state of the Euro-zone business cycle. Indeed our estimates of the business cycle in 2001 will, in all probability, be revised when new data become available. Therefore the most recent estimates about the Euro-zone are subject to change too.

6 Conclusion

Empirical inference about individual Euro-zone business cycles is found to be sensitive to the measure of the business cycle considered. Examining seven measures of the 'growth' business cycle, and three measures of the 'classical' business cycle, we find that business cycles of Euro-zone countries display different properties according to the measure used. Furthermore, correlation between individual Euro-zone countries varies according to the measure considered.

However, our measure of synchronisation between Euro-zone business cycles, i.e. the average (size-weighted) correlation between them, exhibits common features across alternative measures of the business cycle. Average correlation, computed for a series of rolling windows, has risen from zero in the mid 1980s, and the most recent estimates for the turn of the Century indicate that average correlation continues to rise. This increased synchronisation between Euro-zone business cycles is consistent with the emergence of a 'common' Euro-zone business cycle. Perhaps we can only expect this trend to continue since the very existence of EMU, and irrevocably fixed exchange-rates, is believed to encourage the emergence of a 'common' Euro-zone business cycle; for further discussion of such endogeneities see Artis (1999). Accompanying this increased correlation our results indicate that Euro-zone business cycles have moved 'closer' together, according to the (size-weighted) average root mean squared difference computed for a series of rolling windows. Together, the increased average correlation and reduced cyclical disparity bode well for the successful operation of a common monetary policy in Europe.

References

Agresti, A.-M., and B. Mojon (2001): “Some stylised facts on the Euro area business cycle,” ECB Working Paper No. 95.

Arino, M. A., and P. Newbold (1998): “Computation of the Beveridge-Nelson decomposition for multivariate economic time series,” *Economics Letters*, 61(1), 37–42.

Artis, M. (1999): “Should the UK join EMU?,” *National Institute Economic Review*, 171, 70–81.

Artis, M., H.-M. Krolzig, and J. Toro (1999): “The European business cycle,” *CEPR Discussion Paper*, 2242.

Artis, M., and W. Zhang (1997): “International business cycle and the ERM: is there a European business cycle?,” *International Journal of Finance and Economics*, 2, 1–16.

——— (1999): “Further evidence on the international business cycle and the ERM: is there a European business cycle?,” *Oxford Economic Papers*, 51, 120–132.

Backhouse, R., and A. Salanti (eds.) (2000): *Macroeconomics and the Real World.* Oxford University Press.

Baxter, M., and R. G. King (1999): “Measuring business cycles: Approximate band-pass filters for economic time series,” *Review of Economics and Statistics*, 81(4), 575–93.

Benati, L. (2001): “Band pass filtering, cointegration and business cycle analysis,” Bank of England Working Paper no. 142.

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.

Bry, G., and C. Boschan (1971): “Interpretation and analysis of time-series scatters,” *The American Statistician*, 25(2), 29–33.

Burns, A., and W. Mitchell (1946): *Measuring Business Cycles.* NBER, New York.

Canova, F. (1998a): “Detrending and business cycle facts,” *Journal of Monetary Economics*, 41(3), 475–512.

——— (1998b): “Detrending and business cycle facts: A user’s guide,” *Journal of Monetary Economics*, 41(3), 533–40.

——— (1999): “Does detrending matter for the determination of the reference cycle and the selection of turning points?,” *Economic Journal*, 109(452), 126–50.

Chourdakis, K., and E. Tzavalis (2000): “Option pricing under discrete shifts in stocks returns,” *Queen Mary University of London working paper series*, 426.

Christiano, L. J., and T. J. Fitzgerald (1999): “The band pass filter,” NBER Working Paper 7257.

Christodoulakis, N., S. Dimelis, and T. Kollintzas (1995): “Comparison of business cycles in the EC: Idiosyncracies and regularities,” *Economica*, 62, 1–27.

Clark, T., and E. van Wincoop (2001): “Borders and business cycles,” *Journal of International Economics*, 55, 59–85.

Doornik, J. (1998): *Object-oriented Matrix Programming using Ox.* Timberlake Consultants.

Durland, J., and T. McCurdy (1994): “Duration-dependent transitions in a Markov model of US GNP growth,” *Journal of Business and Economic Statistics*, 12, 279–288.

Filardo, A. (1994): “Business-Cycle phases and their transition dynamics,” *Journal of Business and Economic Statistics*, 12, 299–308.

——— (1998): “Choosing information variables for transition probabilities in a time-varying transition probability Markov switching model,” Discussion Paper No. 98-09, Federal Reserve Bank of Kansas City.

Findley, D., B. Monsell, W. Bell, M. Otto, and B.-C. Chen (1998): “New capabilities and methods of the X-12-ARIMA seasonal adjustment program,” *Journal of Business and Economic Statistics*, 16, 127–152.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): “Reference cycles: the NBER methodology revisited,” Discussion Paper No. 2400, CEPR.

Frankel, J., and A. Rose (1998): “The endogenenity of the optimum currency area criteria,” *Economic Journal*, 108, 1009–1025.

Garcia, R., and P. Perron (1996): “An analysis of the real interest rate under regime shifts,” *Review of Economics and Statistics*, 78, 111–125.

Gomez, V., and A. Maravall (1998): “Tramo-Seats,” Mimeo, Banca de Espana.

Gros, D., and N. Thygesen (1998): *European Monetary Integration: from the European Monetary System to Economic and Monetary Union.* Longman, 2nd edition.

Hamilton, J. D. (1989): “A new approach to the analysis of non-stationary time series and the business cycle,” *Econometrica*, 57, 357–384.

Hansen, B. (1992): “The likelihood ratio test under non-standard conditions: testing the Markov switching model of GNP,” *Journal of Applied Econometrics*, 7, 561–582.

——— (1996): “Erratum: The likelihood ratio test under non-standard conditions: testing the Markov switching model of GNP,” *Journal of Applied Econometrics*, 11, 195–198.

Harding, D., and A. Pagan (2000): “Synchronisation of cycles,” Mimeo, University of Melbourne.

——— (2001): “Extracting, analysing and using cyclical information,” Mimeo, University of Melbourne.

——— (2002a): “A comparison of two business cycle dating methods,” *Journal of Economic Dynamics and Control*, Forthcoming.

——— (2002b): "Dissecting the cycle: A methodological investigation," *Journal of Monetary Economics*, 49, 365–381.

Harvey, A., and S. Koopman (2000): "Signal extraction and the formulation of unobserved components models," *Econometrics Journal*, 3, 84–107.

Harvey, A., and T. Trimbur (2002): "General model based filters for extracting trends and cycles in economic time series," *Review of Economics and Statistics*, Forthcoming.

Harvey, A. C. (1993): *Time Series Models*. Harvester-Wheatsheaf, 2nd. ed.

Harvey, A. C., and A. Jaeger (1993): "Detrending, stylized facts and the business cycle," *Journal of Applied Econometrics*, 8, 231–247.

Inklaar, R., and J. de Haan (2001): "Is there really a European business cycle? A comment," *Oxford Economic Papers*, 53, 215–220.

Kim, C. J. (1994): "Dynamic linear models with Markov-switching," *Journal of Econometrics*, 60(1-2), 1–22.

Koopman, S. J., N. Shephard, and J. A. Doornik (1999): "Statistical algorithms for models in state space using SsfPack 2.2," *Econometrics Journal*, 2(1), 107–60.

Lam, P. S. (1990): "The Hamilton model with a general autoregressive component: Estimation and comparison with other models of economic time series," *Journal of Monetary Economics*, 26(3), 409–32.

Massmann, M., and J. Mitchell (2002a): "Have UK and Eurozone business cycles become more correlated?," *National Institute Economic Review*, 18, 58–71.

——— (2002b): "Re-considering the evidence: are European business cycles converging?," Mimeo, National Institute of Economic and Social Research.

Massmann, M., J. Mitchell, and M. Weale (2002): *Business Cycle Analysis*. Final Report submitted to EUROSTAT by the National Institute of Economic and Social Research.

——— (2003): "Business cycles and turning points: a survey of statistical techniques," *National Institute Economic Review*, Forthcoming.

Morley, J., C. R. Nelson, and E. Zivot (2002): "Why are Beveridge-Nelson and unobserved component decompositions of GDP so different?," *Review of Economics and Statistics*, Forthcoming.

Murray, C. J. (2001): "Cyclical properties of Baxter-King filtered time series," Mimeo, University of Houston.

Newbold, P. (1990): "Precise and efficient computation of the Beveridge-Nelson decomposition of economic time series," *Journal of Monetary Economics*, 26(3), 453–57.

Newey, W., and K. West (1987): "A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

Osborn, D. R. (1995): "Moving average detrending and the analysis of business cycles," *Oxford Bulletin of Economics and Statistics*, 57, 547–558.

Pedersen, T. M. (2001): "The Hodrick-Prescott filter, the Slutzky effect and the distortionary effect of filters," *Journal of Economic Dynamics and Control*, 25, 1081–1101.

Psaradakis, Z., and N. Spagnolo (2003): "On the determination of the number of regimes in Markov-Switching autoregressive models," *Journal of Time Series Analysis*, Forthcoming.

Ravn, M., and H. Uhlig (2002): "On adjusting the Hodrick-Prescott filters for the frequency of observations," *Review of Economics and Statistics*, 84, 371–380.

Vahid, F., and R. Engle (1993): "Common trends and common cycles," *Journal of Applied Econometrics*, 8, 341–360.

Zarnowitz, V., and A. Ozyildirim (2002): "Time series decomposition and measurement of business cycles, trends and growth cycles," NBER Working Paper 8736.

Table 1: Business cycle characteristics: Austria

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.026	-0.180	-0.174	57.059	10.000	11.000	23.200	23.700	51.134	0.459	0.080
TIM	0.077	0.087	-0.777	∞	9.000	10.000	22.556	29.556	57.113	0.423	0.089
MA	0.037	0.084	-0.603	86.444	6.000	7.000	26.500	31.000	52.442	0.508	0.089
HP	0.028	-0.210	-0.004	64.667	10.000	11.000	21.700	24.500	52.784	0.497	0.078
BK	0.019	-0.070	-0.150	55.067	9.000	8.000	22.250	26.750	53.995	0.429	0.089
CF	0.024	-0.163	-0.351	55.067	9.000	8.000	22.500	25.750	53.511	0.424	0.092
PAT	0.044	0.812	1.310	∞	9.000	9.000	28.000	25.000	47.269	0.471	0.075

Notes: UC: unobserved components; TIM: time trend; MA: moving average; HP: Hodrick-Prescott; BK: Baxter-King; CF: Christiano-Fitzgerald; PAT: Phase Average Trend; s.d. denotes standard deviation of cycle; skew: measure of skewness, equal to zero for symmetrical distributions; kurtosis: measure of kurtosis in excess of 3, equal to zero for the normal distribution; period denotes the length in months of the cycle derived from that frequency where the spectral density of the cycle has a peak; 8 indicates that the peak was at frequency zero; peaks and troughs are the number of peaks and troughs identified using the Harding-Pagan rule; PT denotes peak to trough; TP denotes trough to peak; dur denotes the duration of a phase; PE is the percentage of time spent in an expansion state C is the measure of coherence and se(C) its robust estimated standard error. Analysis of the UC, TIM and HP cycles is over the full sample period, 1962m1-2002m5. The BK and CF cycles are considered over the periods 1965m1-1999m5 and 1964m1-2000m5 respectively, and the PAT and MA cycles over the period 1962m1-2001m8 and 1966m1-1998m5, respectively

Table 2: Business cycle characteristics: Belgium

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.038	0.091	-0.401	∞	11.000	11.000	20.700	23.364	52.990	0.482	0.086
TIM	0.089	0.153	-0.391	∞	10.000	10.000	17.333	30.800	63.505	0.636	0.090
MA	0.040	0.145	-0.285	389.000	9.000	9.000	18.125	22.778	52.699	0.462	0.096
HP	0.031	-0.124	0.177	64.667	11.000	11.000	19.200	24.727	56.082	0.523	0.087
BK	0.020	0.094	-0.047	48.588	13.000	13.000	12.083	16.538	52.058	0.155	0.093
CF	0.023	-0.156	-0.085	43.474	12.000	12.000	13.545	17.333	50.363	0.101	0.093

Table 3: Business cycle characteristics: Finland

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.052	-0.405	0.073	∞	12.000	11.000	21.000	21.000	48.454	0.456	0.078
TIM	0.097	-0.394	-1.042	∞	14.000	14.000	12.692	21.000	60.619	0.552	0.093
MA	0.043	-0.483	-0.185	∞	10.000	9.000	15.333	26.000	62.725	0.549	0.114
HP	0.032	-0.236	-0.129	69.286	12.000	11.000	14.273	24.273	63.711	0.520	0.100
BK	0.021	-0.491	-0.228	63.538	13.000	12.000	13.917	18.917	55.932	0.407	0.094
CF	0.026	0.068	-0.202	59.000	11.000	10.000	17.100	21.600	53.269	0.408	0.090
PAT	0.073	0.103	1.851	∞	11.000	10.000	20.000	26.200	55.882	0.507	0.088

Table 4: Business cycle characteristics: France

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.017	-0.387	0.758	40.417	13.000	13.000	19.500	15.923	42.680	0.405	0.076
TIM	0.092	-0.152	-0.888	∞	11.000	11.000	22.400	19.727	44.742	0.412	0.079
MA	0.024	-0.083	-0.285	259.333	9.000	9.000	20.556	21.000	52.442	0.465	0.092
HP	0.019	-0.372	0.600	51.053	14.000	14.000	17.154	15.571	44.948	0.334	0.080
BK	0.014	-0.398	0.609	45.889	12.000	11.000	16.818	17.909	51.816	0.384	0.085
CF	0.016	-0.611	0.680	45.889	11.000	10.000	18.500	19.700	51.816	0.352	0.087
PAT	0.039	-2.391	20.921	∞	11.000	12.000	17.636	21.364	56.092	0.711	0.075

Table 5: Business cycle characteristics: Germany

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.035	-0.412	0.074	74.615	12.000	11.000	16.091	25.909	60.412	0.559	0.086
TIM	0.081	-0.325	-0.822	∞	10.000	10.000	18.333	30.100	62.062	0.744	0.081
MA	0.038	-0.309	-0.140	77.800	6.000	7.000	21.500	35.333	62.468	0.590	0.097
HP	0.029	-0.532	0.325	60.625	12.000	11.000	16.727	24.727	57.732	0.519	0.085
BK	0.021	-0.532	0.178	55.067	11.000	11.000	18.000	19.455	51.816	0.462	0.083
CF	0.028	-0.453	-0.135	55.067	10.000	9.000	19.889	23.222	53.269	0.427	0.088
PAT	0.040	-0.261	-0.382	86.545	12.000	11.000	15.727	26.273	62.395	0.741	0.085

Table 6: Business cycle characteristics: Greece

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.028	0.351	0.463	53.889	10.000	11.000	19.800	23.900	53.196	0.378	0.084
TIM	0.221	-0.130	-1.273	∞	8.000	8.000	30.250	30.000	50.103	0.349	0.083
MA	0.036	0.908	1.181	778.000	8.000	8.000	24.625	25.571	49.357	0.359	0.091
HP	0.029	0.430	0.680	57.059	10.000	11.000	19.800	23.900	53.196	0.378	0.084
BK	0.019	1.119	2.210	45.889	11.000	12.000	17.000	17.727	52.542	0.184	0.090
CF	0.023	0.931	1.672	51.625	11.000	12.000	17.091	17.636	52.542	0.170	0.087
PAT	0.042	0.406	0.654	∞	9.000	9.000	17.889	35.500	66.176	0.646	0.106

Table 7: Business cycle characteristics: Ireland

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.039	-0.395	0.229	64.667	14.000	14.000	18.643	16.462	46.186	0.297	0.081
TIM	0.173	0.744	0.392	∞	13.000	13.000	17.538	15.417	52.990	0.269	0.090
MA	0.046	-0.178	-0.418	77.800	10.000	11.000	19.400	17.000	48.843	0.354	0.095
HP	0.037	-0.301	0.384	57.059	13.000	13.000	19.077	18.917	48.866	0.322	0.083
BK	0.025	-0.219	-0.375	51.625	12.000	12.000	16.000	17.667	51.332	0.403	0.089
CF	0.031	0.155	0.057	51.625	10.000	10.000	20.333	20.400	49.395	0.372	0.087
PAT	0.105	1.598	3.639	∞	13.000	12.000	17.000	21.833	55.882	0.480	0.091

Table 8: Business cycle characteristics: Italy

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.024	0.065	0.243	31.290	14.000	15.000	16.929	16.857	49.897	0.297	0.086
TIM	0.077	-0.410	0.568	∞	9.000	9.000	20.625	33.000	61.237	0.647	0.087
MA	0.036	-0.176	-0.543	778.000	9.000	9.000	17.000	26.889	62.211	0.689	0.100
HP	0.029	-0.196	-0.253	32.333	13.000	13.000	18.500	18.462	49.485	0.402	0.083
BK	0.021	-0.184	0.006	31.769	14.000	13.000	13.154	15.308	54.479	0.432	0.088
CF	0.024	0.261	0.625	31.769	13.000	12.000	14.083	16.833	54.964	0.296	0.090
PAT	0.041	-0.023	-0.153	68.000	11.000	11.000	19.200	24.545	56.723	0.708	0.078

Table 9: Business cycle characteristics: Luxembourg

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.048	-0.075	-0.380	88.182	13.000	12.000	19.167	16.917	49.691	0.579	0.071
TIM	0.054	-0.043	-0.288	∞	12.000	11.000	20.273	17.364	47.216	0.546	0.072
MA	0.045	0.140	-0.253	64.833	10.000	11.000	18.200	18.700	52.185	0.606	0.080
HP	0.039	0.006	-0.085	57.059	12.000	11.000	19.545	18.091	48.866	0.506	0.076
BK	0.025	0.066	-0.512	51.625	11.000	12.000	17.909	15.909	43.826	0.453	0.071
CF	0.032	0.351	0.000	55.067	11.000	12.000	18.455	15.545	42.857	0.502	0.067
PAT	0.059	0.010	0.034	68.000	12.000	12.000	16.727	22.083	55.672	0.748	0.070

Table 10: Business cycle characteristics: The Netherlands

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.023	-0.315	0.052	51.053	14.000	13.000	16.308	19.615	54.021	0.489	0.083
TIM	0.080	-0.420	0.134	∞	12.000	12.000	17.833	23.364	55.876	0.670	0.075
MA	0.030	-0.320	-0.012	86.444	8.000	9.000	22.750	21.750	48.843	0.591	0.081
HP	0.025	-0.414	0.202	64.667	12.000	13.000	19.000	19.667	50.103	0.528	0.079
BK	0.016	-0.349	-0.223	51.625	11.000	11.000	16.200	18.727	49.879	0.450	0.079
CF	0.019	-0.139	-0.868	48.588	10.000	10.000	19.444	19.200	46.489	0.511	0.074
PAT	0.031	-0.399	0.301	86.545	12.000	11.000	20.273	21.636	51.471	0.435	0.086

Table 11: Business cycle characteristics: Portugal

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.045	-0.054	0.225	74.615	10.000	11.000	28.100	17.000	38.351	0.337	0.075
TIM	0.073	0.137	-1.129	∞	12.000	13.000	21.250	16.333	43.711	0.368	0.078
MA	0.049	-0.016	0.028	70.727	7.000	8.000	22.857	20.429	41.131	0.474	0.078
HP	0.039	0.037	0.486	64.667	11.000	12.000	22.818	18.182	44.536	0.296	0.084
BK	0.026	0.146	0.377	55.067	12.000	13.000	17.417	16.000	48.184	0.411	0.080
CF	0.032	0.224	0.311	55.067	12.000	12.000	18.273	16.250	47.215	0.355	0.083
PAT	0.052	0.025	0.376	95.200	13.000	14.000	20.538	14.154	40.126	0.419	0.070

Table 12: Business cycle characteristics: Spain

	Statistical Properties				Turning point rule						
	s.d.	skew	kurtosis	period	peaks	troughs	dur:PT	dur:TP	PE	C	se(C)
UC	0.023	-0.133	-0.161	42.174	11.000	11.000	23.500	18.545	42.062	0.203	0.085
TIM	0.184	-0.388	-0.278	∞	10.000	10.000	26.000	22.000	45.361	0.382	0.078
MA	0.037	-0.005	-0.068	∞	8.000	8.000	21.625	25.857	55.527	0.477	0.095
HP	0.028	-0.234	-0.258	53.889	11.000	10.000	25.800	18.800	41.237	0.279	0.082
BK	0.020	-0.342	0.254	48.588	12.000	11.000	18.000	16.818	48.910	0.243	0.091
CF	0.024	-0.070	0.296	48.588	12.000	11.000	16.727	18.000	52.058	0.283	0.091
PAT	0.037	-0.007	-0.223	∞	11.000	11.000	19.700	22.000	50.840	0.437	0.083

Table 13: Mean business cycle duration for three measures of the 'classical' business cycle

	Harding-Pagan				MS: constant tran. prob.				MS: time-varying	
	1962m1-2002m5		1970m1-2002m5		C_u	E_u	C_b	E_b	C_b	E_b
	PT	TP	PT	TP	PT	TP	PT	TP	PT	TP
AUS	13.63	45.00	13.67	45.00	14.92	20.40	8.58	11.50	11.35	11.06
BEL	19.33	49.71	19.33	32.86	14.49	8.33	12.90	6.58	13.945	5.53
FIN	11.33	54.86	13.25	58.80	12.20	33.33	8.08	21.00	7.33	21.69
FRA	13.55	33.80	12.75	27.75	33.33	32.25	20.44	18.60	18.43	8.00
GER	13.11	34.80	13.00	27.33	12.82	23.25	13.10	23.90	10.46	18.00
GRE	10.50	51.57	10.50	36.60	27.78	13.88	22.73	10.919	16.73	7.93
IRE	17.75	78.20	19.00	89.00	14.92	15.87	9.211	10.83	7.95	11.11
ITA	16.13	43.00	17.43	33.43	13.51	15.51	10.93	12.88	11.40	12.44
LUX	11.83	25.46	13.38	27.67	17.86	5.55	13.35	4.91	8.50	3.71
NET	13.60	33.80	13.89	25.89	14.08	23.80	9.43	17.00	10.33	14.33
POR	11.89	42.00	12.50	42.83	20.83	12.50	12.61	7.94	13.56	9.56
SPA	23.60	41.20	23.60	41.20	21.73	13.88	13.81	8.77	11.40	6.76

Table 14: UC: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1962.1-2002.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.491	0.283	0.500	0.581	0.199	0.322	0.303	0.333	0.566	0.371	0.391
	-	(0.087)	(0.083)	(0.099)	(0.090)	(0.082)	(0.098)	(0.090)	(0.081)	(0.089)	(0.095)	(0.078)
BEL	-	1.00	0.308	0.494	0.666	0.294	0.454	0.301	0.444	0.493	0.488	0.250
	-	-	(0.093)	(0.085)	(0.090)	(0.080)	(0.089)	(0.083)	(0.079)	(0.083)	(0.101)	(0.076)
FIN	-	-	1.00	0.190	0.295	0.203	0.224	0.105	0.253	0.211	0.197	0.174
	-	-	-	(0.080)	(0.116)	(0.079)	(0.069)	(0.075)	(0.085)	(0.076)	(0.092)	(0.094)
FRA	-	-	-	1.00	0.450	0.210	0.350	0.415	0.307	0.460	0.268	0.449
	-	-	-	-	(0.097)	(0.089)	(0.108)	(0.105)	(0.062)	(0.098)	(0.090)	(0.091)
GER	-	-	-	-	1.00	0.231	0.359	0.167	0.501	0.562	0.430	0.355
	-	-	-	-	-	(0.079)	(0.085)	(0.078)	(0.096)	(0.086)	(0.087)	(0.092)
GRE	-	-	-	-	-	1.00	0.293	0.083	0.160	0.124	0.218	0.286
	-	-	-	-	-	-	(0.088)	(0.086)	(0.076)	(0.076)	(0.092)	(0.076)
IRE	-	-	-	-	-	-	1.00	0.229	0.305	0.376	0.350	0.226
	-	-	-	-	-	-	-	(0.090)	(0.073)	(0.089)	(0.095)	(0.077)
ITA	-	-	-	-	-	-	-	1.00	0.146	0.361	0.081	0.279
	-	-	-	-	-	-	-	-	(0.067)	(0.084)	(0.087)	(0.087)
LUX	-	-	-	-	-	-	-	-	1.00	0.362	0.244	0.233
	-	-	-	-	-	-	-	-	-	(0.072)	(0.082)	(0.077)
NET	-	-	-	-	-	-	-	-	-	1.00	0.346	0.250
	-	-	-	-	-	-	-	-	-	-	(0.086)	(0.074)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.257
	-	-	-	-	-	-	-	-	-	-	-	(0.085)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 15: TIM: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1962.1-2002.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.889	0.809	0.846	0.871	0.734	0.028	0.684	0.379	0.849	0.345	0.836
	-	(0.112)	(0.093)	(0.103)	(0.103)	(0.096)	(0.065)	(0.122)	(0.082)	(0.118)	(0.085)	(0.113)
BEL	-	1.00	0.683	0.778	0.810	0.630	0.122	0.777	0.268	0.881	0.221	0.793
	-	-	(0.099)	(0.117)	(0.111)	(0.109)	(0.063)	(0.128)	(0.076)	(0.129)	(0.088)	(0.127)
FIN	-	-	1.00	0.800	0.703	0.754	-0.058	0.546	0.330	0.748	0.389	0.793
	-	-	-	(0.083)	(0.094)	(0.077)	(0.079)	(0.109)	(0.078)	(0.099)	(0.085)	(0.092)
FRA	-	-	-	1.00	0.916	0.955	-0.391	0.619	0.453	0.879	0.582	0.969
	-	-	-	-	(0.102)	(0.091)	(0.062)	(0.126)	(0.069)	(0.121)	(0.069)	(0.112)
GER	-	-	-	-	1.00	0.847	-0.322	0.670	0.508	0.889	0.550	0.900
	-	-	-	-	-	(0.093)	(0.064)	(0.122)	(0.078)	(0.120)	(0.077)	(0.113)
GRE	-	-	-	-	-	1.00	-0.557	0.496	0.432	0.791	0.656	0.942
	-	-	-	-	-	-	(0.074)	(0.118)	(0.063)	(0.111)	(0.068)	(0.102)
IRE	-	-	-	-	-	-	1.00	0.053	-0.306	-0.161	-0.611	-0.375
	-	-	-	-	-	-	-	(0.069)	(0.073)	(0.073)	(0.098)	(0.065)
ITA	-	-	-	-	-	-	-	1.00	0.068	0.815	0.125	0.687
	-	-	-	-	-	-	-	-	(0.076)	(0.142)	(0.087)	(0.141)
LUX	-	-	-	-	-	-	-	-	1.00	0.326	0.513	0.381
	-	-	-	-	-	-	-	-	-	(0.067)	(0.076)	(0.062)
NET	-	-	-	-	-	-	-	-	-	1.00	0.382	0.907
	-	-	-	-	-	-	-	-	-	-	(0.086)	(0.136)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.536
	-	-	-	-	-	-	-	-	-	-	-	(0.071)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 16: MA: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1966.1-1998.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.677	0.447	0.674	0.765	0.509	0.478	0.378	0.449	0.610	0.444	0.530
	-	(0.103)	(0.100)	(0.103)	(0.104)	(0.110)	(0.102)	(0.104)	(0.098)	(0.103)	(0.116)	(0.108)
BEL	-	1.00	0.381	0.723	0.660	0.501	0.568	0.493	0.425	0.670	0.502	0.581
	-	-	(0.094)	(0.105)	(0.094)	(0.106)	(0.106)	(0.104)	(0.081)	(0.105)	(0.114)	(0.107)
FIN	-	-	1.00	0.400	0.281	0.380	0.330	0.395	0.245	0.367	0.181	0.452
	-	-	-	(0.099)	(0.118)	(0.100)	(0.083)	(0.096)	(0.091)	(0.090)	(0.104)	(0.125)
FRA	-	-	-	1.00	0.673	0.542	0.574	0.412	0.472	0.588	0.436	0.665
	-	-	-	-	(0.106)	(0.108)	(0.110)	(0.111)	(0.079)	(0.110)	(0.107)	(0.116)
GER	-	-	-	-	1.00	0.373	0.419	0.301	0.510	0.698	0.401	0.448
	-	-	-	-	-	(0.096)	(0.090)	(0.097)	(0.105)	(0.108)	(0.100)	(0.094)
GRE	-	-	-	-	-	1.00	0.404	0.294	0.200	0.335	0.357	0.595
	-	-	-	-	-	-	(0.116)	(0.094)	(0.093)	(0.092)	(0.119)	(0.135)
IRE	-	-	-	-	-	-	1.00	0.509	0.395	0.565	0.406	0.441
	-	-	-	-	-	-	-	(0.105)	(0.076)	(0.108)	(0.112)	(0.106)
ITA	-	-	-	-	-	-	-	1.00	0.195	0.548	0.278	0.454
	-	-	-	-	-	-	-	-	(0.089)	(0.115)	(0.107)	(0.116)
LUX	-	-	-	-	-	-	-	-	1.00	0.467	0.267	0.338
	-	-	-	-	-	-	-	-	-	(0.090)	(0.085)	(0.090)
NET	-	-	-	-	-	-	-	-	-	1.00	0.388	0.464
	-	-	-	-	-	-	-	-	-	-	(0.108)	(0.114)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.485
	-	-	-	-	-	-	-	-	-	-	-	(0.117)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 17: HP: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1962.1-2002.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.582	0.462	0.558	0.681	0.318	0.373	0.334	0.411	0.542	0.373	0.469
	-	(0.095)	(0.086)	(0.106)	(0.100)	(0.084)	(0.102)	(0.099)	(0.075)	(0.093)	(0.100)	(0.086)
BEL	-	1.00	0.387	0.627	0.595	0.321	0.451	0.428	0.409	0.593	0.413	0.434
	-	-	(0.079)	(0.104)	(0.089)	(0.078)	(0.093)	(0.100)	(0.057)	(0.090)	(0.096)	(0.087)
FIN	-	-	1.00	0.347	0.349	0.236	0.227	0.338	0.296	0.364	0.156	0.406
	-	-	-	(0.080)	(0.096)	(0.072)	(0.069)	(0.078)	(0.069)	(0.075)	(0.083)	(0.092)
FRA	-	-	-	1.00	0.577	0.335	0.432	0.430	0.395	0.512	0.351	0.542
	-	-	-	-	(0.107)	(0.088)	(0.106)	(0.110)	(0.060)	(0.105)	(0.089)	(0.098)
GER	-	-	-	-	1.00	0.250	0.321	0.226	0.502	0.660	0.383	0.451
	-	-	-	-	-	(0.083)	(0.082)	(0.095)	(0.085)	(0.096)	(0.083)	(0.094)
GRE	-	-	-	-	-	1.00	0.275	0.254	0.133	0.238	0.236	0.428
	-	-	-	-	-	-	(0.091)	(0.080)	(0.077)	(0.070)	(0.100)	(0.092)
IRE	-	-	-	-	-	-	1.00	0.384	0.337	0.451	0.334	0.310
	-	-	-	-	-	-	-	(0.096)	(0.069)	(0.092)	(0.094)	(0.083)
ITA	-	-	-	-	-	-	-	1.00	0.168	0.429	0.177	0.380
	-	-	-	-	-	-	-	-	(0.070)	(0.098)	(0.097)	(0.097)
LUX	-	-	-	-	-	-	-	-	1.00	0.456	0.233	0.300
	-	-	-	-	-	-	-	-	-	(0.070)	(0.065)	(0.069)
NET	-	-	-	-	-	-	-	-	-	1.00	0.352	0.359
	-	-	-	-	-	-	-	-	-	-	(0.085)	(0.088)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.426
	-	-	-	-	-	-	-	-	-	-	-	(0.094)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 18: BK: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1965.1-1999.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.767	0.566	0.705	0.800	0.428	0.526	0.442	0.580	0.748	0.489	0.626
	-	(0.140)	(0.114)	(0.132)	(0.116)	(0.109)	(0.131)	(0.142)	(0.104)	(0.130)	(0.141)	(0.120)
BEL	-	1.00	0.450	0.794	0.694	0.455	0.613	0.509	0.559	0.747	0.539	0.634
	-	-	(0.117)	(0.143)	(0.109)	(0.103)	(0.130)	(0.142)	(0.087)	(0.133)	(0.132)	(0.122)
FIN	-	-	1.00	0.393	0.359	0.270	0.373	0.358	0.382	0.407	0.159	0.477
	-	-	-	(0.099)	(0.118)	(0.092)	(0.093)	(0.106)	(0.095)	(0.099)	(0.118)	(0.124)
FRA	-	-	-	1.00	0.652	0.533	0.573	0.519	0.551	0.680	0.409	0.728
	-	-	-	-	(0.123)	(0.111)	(0.139)	(0.141)	(0.086)	(0.141)	(0.119)	(0.124)
GER	-	-	-	-	1.00	0.403	0.415	0.294	0.678	0.787	0.455	0.591
	-	-	-	-	-	(0.113)	(0.103)	(0.116)	(0.120)	(0.126)	(0.112)	(0.120)
GRE	-	-	-	-	-	1.00	0.402	0.348	0.272	0.370	0.293	0.568
	-	-	-	-	-	-	(0.137)	(0.113)	(0.109)	(0.100)	(0.150)	(0.133)
IRE	-	-	-	-	-	-	1.00	0.510	0.554	0.574	0.434	0.430
	-	-	-	-	-	-	-	(0.132)	(0.096)	(0.128)	(0.133)	(0.114)
ITA	-	-	-	-	-	-	-	1.00	0.356	0.561	0.232	0.462
	-	-	-	-	-	-	-	-	(0.105)	(0.142)	(0.147)	(0.136)
LUX	-	-	-	-	-	-	-	-	1.00	0.684	0.309	0.495
	-	-	-	-	-	-	-	-	-	(0.108)	(0.098)	(0.103)
NET	-	-	-	-	-	-	-	-	-	1.00	0.443	0.491
	-	-	-	-	-	-	-	-	-	-	(0.129)	(0.129)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.500
	-	-	-	-	-	-	-	-	-	-	-	(0.132)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 19: CF: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1964.1-2000.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.762	0.455	0.663	0.800	0.381	0.533	0.387	0.557	0.818	0.459	0.592
	-	(0.131)	(0.098)	(0.135)	(0.108)	(0.105)	(0.136)	(0.133)	(0.088)	(0.119)	(0.123)	(0.114)
BEL	-	1.00	0.355	0.774	0.662	0.387	0.646	0.449	0.596	0.765	0.562	0.597
	-	-	(0.105)	(0.145)	(0.096)	(0.105)	(0.131)	(0.144)	(0.074)	(0.118)	(0.130)	(0.119)
FIN	-	-	1.00	0.258	0.204	0.231	0.279	0.400	0.336	0.278	0.044	0.489
	-	-	-	(0.093)	(0.099)	(0.074)	(0.092)	(0.112)	(0.072)	(0.090)	(0.118)	(0.109)
FRA	-	-	-	1.00	0.614	0.442	0.602	0.436	0.589	0.657	0.404	0.672
	-	-	-	-	(0.106)	(0.112)	(0.139)	(0.125)	(0.086)	(0.123)	(0.111)	(0.123)
GER	-	-	-	-	1.00	0.255	0.359	0.116	0.677	0.817	0.369	0.487
	-	-	-	-	-	(0.106)	(0.092)	(0.101)	(0.117)	(0.110)	(0.088)	(0.109)
GRE	-	-	-	-	-	1.00	0.415	0.378	0.122	0.285	0.288	0.591
	-	-	-	-	-	-	(0.141)	(0.126)	(0.092)	(0.106)	(0.144)	(0.134)
IRE	-	-	-	-	-	-	1.00	0.450	0.579	0.539	0.505	0.439
	-	-	-	-	-	-	-	(0.131)	(0.101)	(0.121)	(0.134)	(0.117)
ITA	-	-	-	-	-	-	-	1.00	0.309	0.432	0.249	0.483
	-	-	-	-	-	-	-	-	(0.091)	(0.126)	(0.161)	(0.134)
LUX	-	-	-	-	-	-	-	-	1.00	0.756	0.304	0.403
	-	-	-	-	-	-	-	-	-	(0.101)	(0.084)	(0.089)
NET	-	-	-	-	-	-	-	-	-	1.00	0.459	0.420
	-	-	-	-	-	-	-	-	-	-	(0.115)	(0.103)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.516
	-	-	-	-	-	-	-	-	-	-	-	(0.136)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 20: PAT: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1962.1-2001.8

	AUT	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.591	0.666	0.636	0.621	0.644	0.351	0.467	0.576	0.315	0.351
	-	(0.156)	(0.101)	(0.086)	(0.128)	(0.185)	(0.084)	(0.083)	(0.094)	(0.079)	(0.070)
FIN	-	1.00	0.459	0.159	0.551	0.753	0.175	0.188	0.299	0.116	0.150
	-	-	(0.102)	(0.109)	(0.131)	(0.185)	(0.069)	(0.085)	(0.090)	(0.093)	(0.084)
FRA	-	-	1.00	0.581	0.573	0.501	0.460	0.482	0.561	0.374	0.501
	-	-	-	(0.089)	(0.089)	(0.112)	(0.094)	(0.087)	(0.091)	(0.086)	(0.077)
GER	-	-	-	1.00	0.388	0.165	0.371	0.592	0.585	0.514	0.388
	-	-	-	-	(0.081)	(0.098)	(0.098)	(0.082)	(0.097)	(0.105)	(0.076)
GRE	-	-	-	-	1.00	0.601	0.334	0.297	0.394	0.202	0.343
	-	-	-	-	-	(0.158)	(0.074)	(0.085)	(0.082)	(0.080)	(0.079)
IRE	-	-	-	-	-	1.00	0.131	0.131	0.364	0.032	0.106
	-	-	-	-	-	-	(0.058)	(0.078)	(0.091)	(0.072)	(0.068)
ITA	-	-	-	-	-	-	1.00	0.340	0.488	0.256	0.471
	-	-	-	-	-	-	-	(0.110)	(0.109)	(0.094)	(0.083)
LUX	-	-	-	-	-	-	-	1.00	0.572	0.438	0.403
	-	-	-	-	-	-	-	-	(0.101)	(0.096)	(0.092)
NET	-	-	-	-	-	-	-	-	1.00	0.363	0.514
	-	-	-	-	-	-	-	-	-	(0.095)	(0.084)
POR	-	-	-	-	-	-	-	-	-	1.00	0.380
	-	-	-	-	-	-	-	-	-	-	(0.084)
SPA	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-

Table 21: Harding-Pagan: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1962.1-2002.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	0.267	0.302	0.279	0.321	0.080	0.293	0.371	0.415	0.204	0.104	0.300
	-	(0.096)	(0.105)	(0.093)	(0.092)	(0.086)	(0.107)	(0.101)	(0.093)	(0.094)	(0.089)	(0.104)
BEL	-	1.00	0.129	0.333	0.460	0.160	0.260	0.223	0.182	0.378	0.251	0.240
	-	-	(0.096)	(0.093)	(0.098)	(0.100)	(0.106)	(0.094)	(0.092)	(0.097)	(0.104)	(0.094)
FIN	-	-	1.00	0.280	0.230	0.020	0.582	0.338	0.337	0.088	0.314	0.329
	-	-	-	(0.099)	(0.097)	(0.085)	(0.124)	(0.104)	(0.094)	(0.091)	(0.102)	(0.106)
FRA	-	-	-	1.00	0.522	0.208	0.198	0.241	0.272	0.401	0.168	0.382
	-	-	-	-	(0.090)	(0.101)	(0.097)	(0.094)	(0.085)	(0.092)	(0.092)	(0.097)
GER	-	-	-	-	1.00	0.281	0.283	0.325	0.211	0.438	0.328	0.334
	-	-	-	-	-	(0.105)	(0.100)	(0.094)	(0.087)	(0.096)	(0.101)	(0.097)
GRE	-	-	-	-	-	1.00	-0.045	0.202	0.062	0.045	0.111	0.074
	-	-	-	-	-	-	(0.081)	(0.099)	(0.083)	(0.091)	(0.100)	(0.095)
IRE	-	-	-	-	-	-	1.00	0.288	0.326	0.148	0.229	0.218
	-	-	-	-	-	-	-	(0.102)	(0.098)	(0.094)	(0.105)	(0.103)
ITA	-	-	-	-	-	-	-	1.00	0.120	0.385	0.246	0.404
	-	-	-	-	-	-	-	-	(0.087)	(0.096)	(0.096)	(0.102)
LUX	-	-	-	-	-	-	-	-	1.00	0.181	0.256	0.359
	-	-	-	-	-	-	-	-	-	(0.085)	(0.091)	(0.091)
NET	-	-	-	-	-	-	-	-	-	1.00	0.209	0.313
	-	-	-	-	-	-	-	-	-	-	(0.095)	(0.097)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.350
	-	-	-	-	-	-	-	-	-	-	-	(0.100)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 22: Markov-switching constant transition probability model: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1971.8-2002.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	-0.605	0.424	0.181	0.695	0.195	0.265	0.335	0.352	0.589	0.171	0.413
	-	(0.080)	(0.101)	(0.101)	(0.077)	(0.096)	(0.093)	(0.093)	(0.082)	(0.086)	(0.099)	(0.086)
BEL	-	1.00	-0.303	-0.053	-0.604	-0.163	-0.297	-0.550	-0.263	-0.569	-0.097	-0.372
	-	-	(0.102)	(0.100)	(0.088)	(0.087)	(0.089)	(0.086)	(0.080)	(0.090)	(0.094)	(0.083)
FIN	-	-	1.00	0.334	0.351	-0.005	0.305	0.294	0.292	0.287	-0.069	0.226
	-	-	-	(0.105)	(0.107)	(0.094)	(0.092)	(0.099)	(0.073)	(0.108)	(0.102)	(0.091)
FRA	-	-	-	1.00	0.086	0.366	0.005	0.144	0.157	0.068	0.033	0.380
	-	-	-	-	(0.106)	(0.094)	(0.102)	(0.102)	(0.093)	(0.105)	(0.101)	(0.096)
GER	-	-	-	-	1.00	0.118	0.308	0.402	0.370	0.568	0.132	0.286
	-	-	-	-	-	(0.095)	(0.094)	(0.097)	(0.076)	(0.094)	(0.097)	(0.094)
GRE	-	-	-	-	-	1.00	0.015	0.056	0.173	0.163	0.118	0.264
	-	-	-	-	-	-	(0.098)	(0.095)	(0.097)	(0.091)	(0.099)	(0.104)
IRE	-	-	-	-	-	-	1.00	0.338	0.322	0.259	-0.093	0.199
	-	-	-	-	-	-	-	(0.091)	(0.089)	(0.096)	(0.096)	(0.099)
ITA	-	-	-	-	-	-	-	1.00	0.312	0.547	-0.015	0.353
	-	-	-	-	-	-	-	-	(0.084)	(0.088)	(0.099)	(0.093)
LUX	-	-	-	-	-	-	-	-	1.00	0.412	-0.047	0.292
	-	-	-	-	-	-	-	-	-	(0.075)	(0.093)	(0.094)
NET	-	-	-	-	-	-	-	-	-	1.00	0.132	0.274
	-	-	-	-	-	-	-	-	-	-	(0.100)	(0.092)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.185
	-	-	-	-	-	-	-	-	-	-	-	(0.101)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 23: Markov-switching time varying transition probability model: Matrix of correlation coefficients, with estimated standard errors in parentheses. Computed over 1971.8-2002.5

	AUT	BEL	FIN	FRA	GER	GRE	IRE	ITA	LUX	NET	POR	SPA
AUT	1.00	-0.548	0.407	0.403	0.625	0.171	0.332	0.277	0.308	0.522	0.216	0.361
	-	(0.075)	(0.088)	(0.090)	(0.074)	(0.099)	(0.088)	(0.094)	(0.087)	(0.079)	(0.099)	(0.088)
BEL	-	1.00	-0.326	-0.385	-0.552	-0.166	-0.293	-0.515	-0.286	-0.494	-0.105	-0.327
	-	-	(0.106)	(0.074)	(0.093)	(0.082)	(0.091)	(0.086)	(0.076)	(0.087)	(0.094)	(0.082)
FIN	-	-	1.00	0.322	0.321	0.009	0.273	0.295	0.282	0.294	-0.061	0.235
	-	-	-	(0.078)	(0.104)	(0.092)	(0.095)	(0.098)	(0.072)	(0.100)	(0.105)	(0.088)
FRA	-	-	-	1.00	0.334	0.380	0.167	0.326	0.357	0.370	0.099	0.546
	-	-	-	-	(0.088)	(0.106)	(0.093)	(0.092)	(0.093)	(0.082)	(0.094)	(0.091)
GER	-	-	-	-	1.00	0.113	0.318	0.349	0.360	0.492	0.096	0.297
	-	-	-	-	-	(0.093)	(0.094)	(0.096)	(0.077)	(0.086)	(0.098)	(0.091)
GRE	-	-	-	-	-	1.00	0.002	0.084	0.191	0.208	0.101	0.247
	-	-	-	-	-	-	(0.096)	(0.095)	(0.095)	(0.092)	(0.099)	(0.102)
IRE	-	-	-	-	-	-	1.00	0.321	0.307	0.305	-0.034	0.165
	-	-	-	-	-	-	-	(0.091)	(0.083)	(0.093)	(0.100)	(0.097)
ITA	-	-	-	-	-	-	-	1.00	0.329	0.473	-0.008	0.358
	-	-	-	-	-	-	-	-	(0.084)	(0.084)	(0.099)	(0.091)
LUX	-	-	-	-	-	-	-	-	1.00	0.387	-0.019	0.295
	-	-	-	-	-	-	-	-	-	(0.079)	(0.090)	(0.093)
NET	-	-	-	-	-	-	-	-	-	1.00	0.118	0.210
	-	-	-	-	-	-	-	-	-	-	(0.097)	(0.095)
POR	-	-	-	-	-	-	-	-	-	-	1.00	0.164
	-	-	-	-	-	-	-	-	-	-	-	(0.101)
SPA	-	-	-	-	-	-	-	-	-	-	-	1.00
	-	-	-	-	-	-	-	-	-	-	-	-

Table 24: Mean over 1995m1-2000m2 and 1998m1-2000m2 of the size-weighted average correlation between Euro-zone business cycles and those of Germany and the US for selected measures of the business cycle

	1995m1-2000m2		1998m1-2000m2	
	Germany	US	Germany	US
UC	0.439	0.145	0.488	0.170
TIM	0.454	0.267	0.467	0.342
PAT	0.456	0.351	0.461	0.381
HP	0.472	0.378	0.510	0.387
Markov-switching	0.340	-0.001	0.516	0.021
Markov-switching: time-varying	0.322	0.078	0.494	0.103

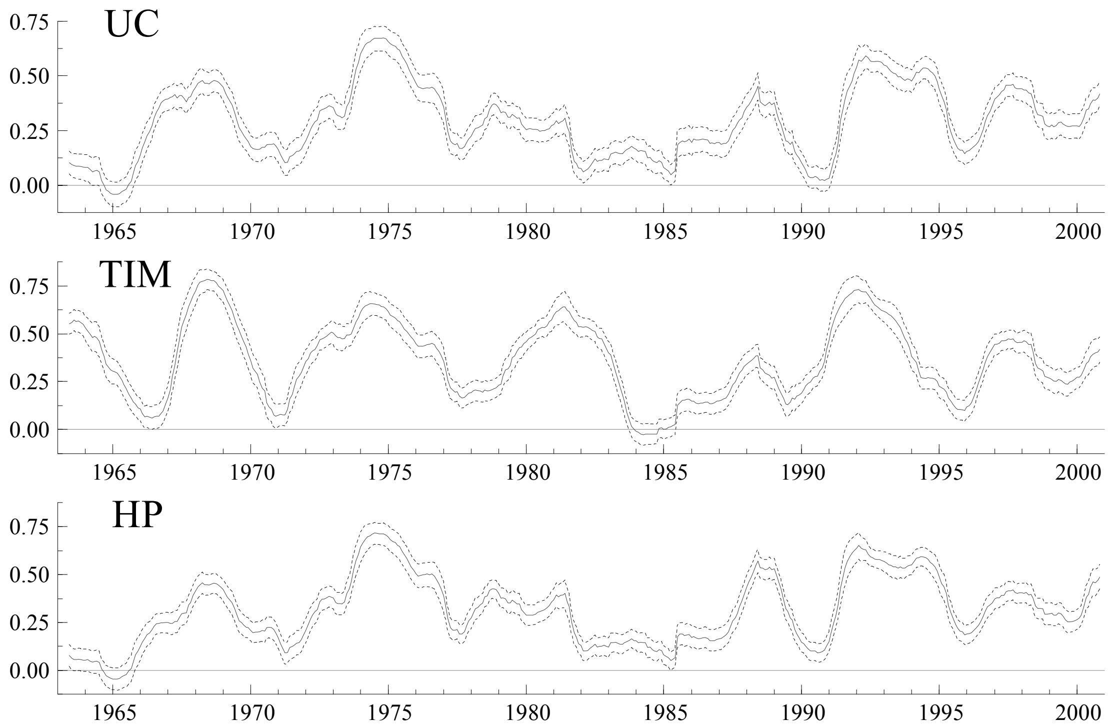

Figure 1: Average correlation between Euro-zone business cycles: UC, TIM and HP growth cycles

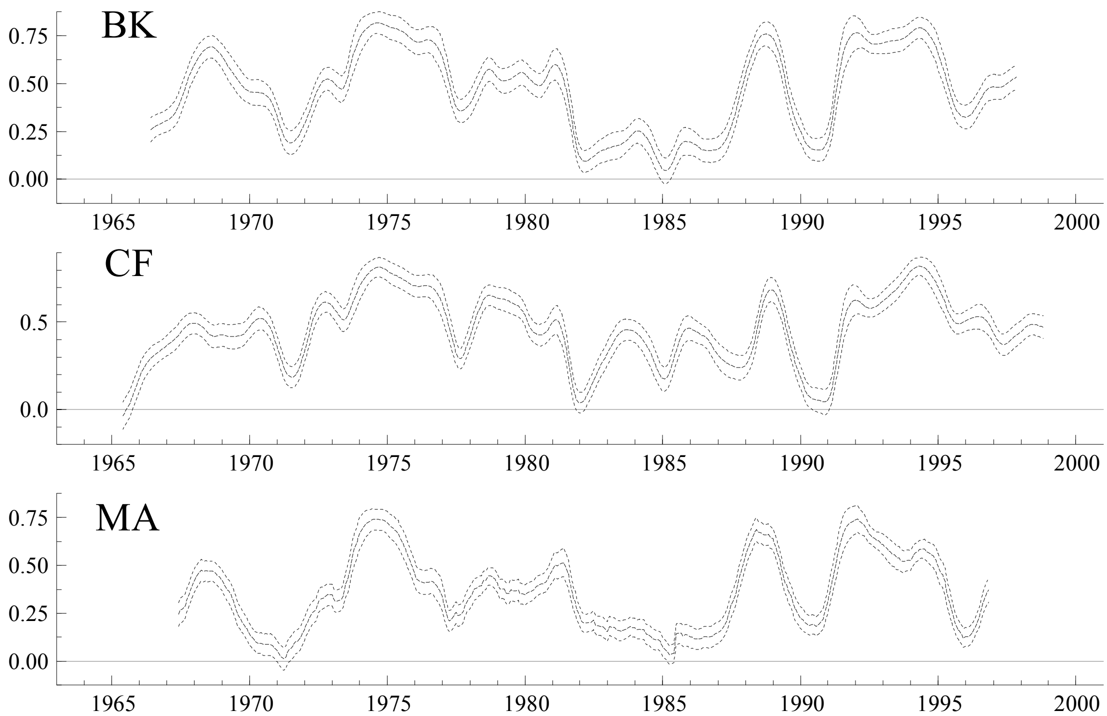

Figure 2: Average correlation between Euro-zone business cycles: BK, CF and MA growth cycles

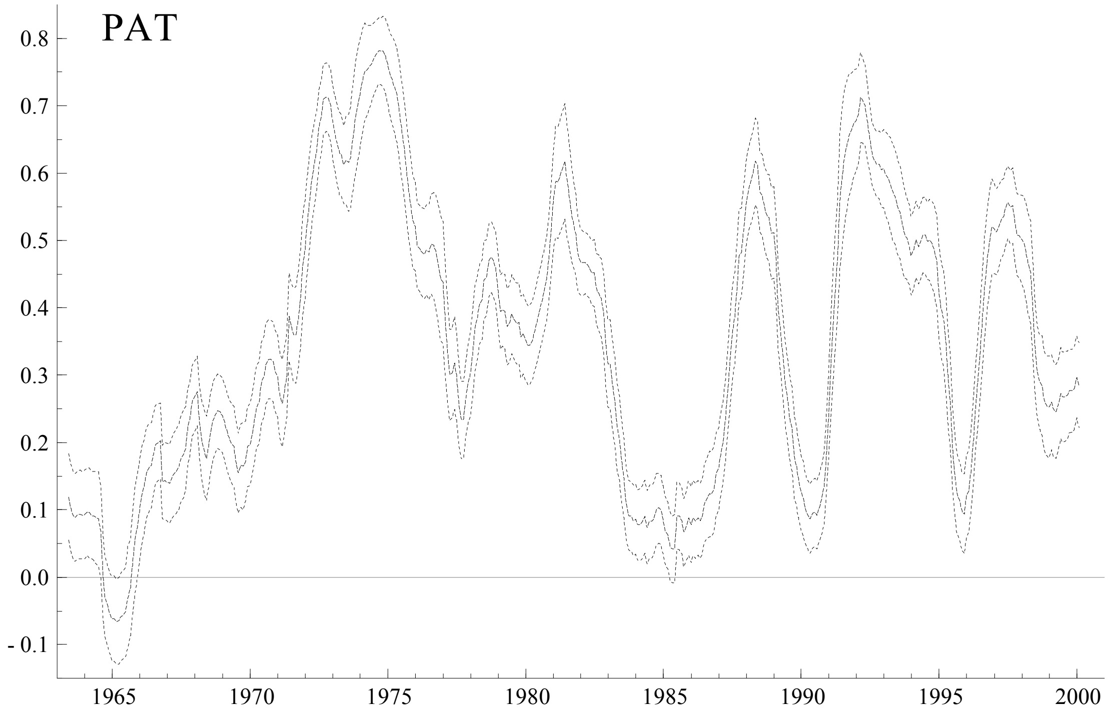

Figure 3: Average correlation between Euro-zone business cycles: PAT growth cycle

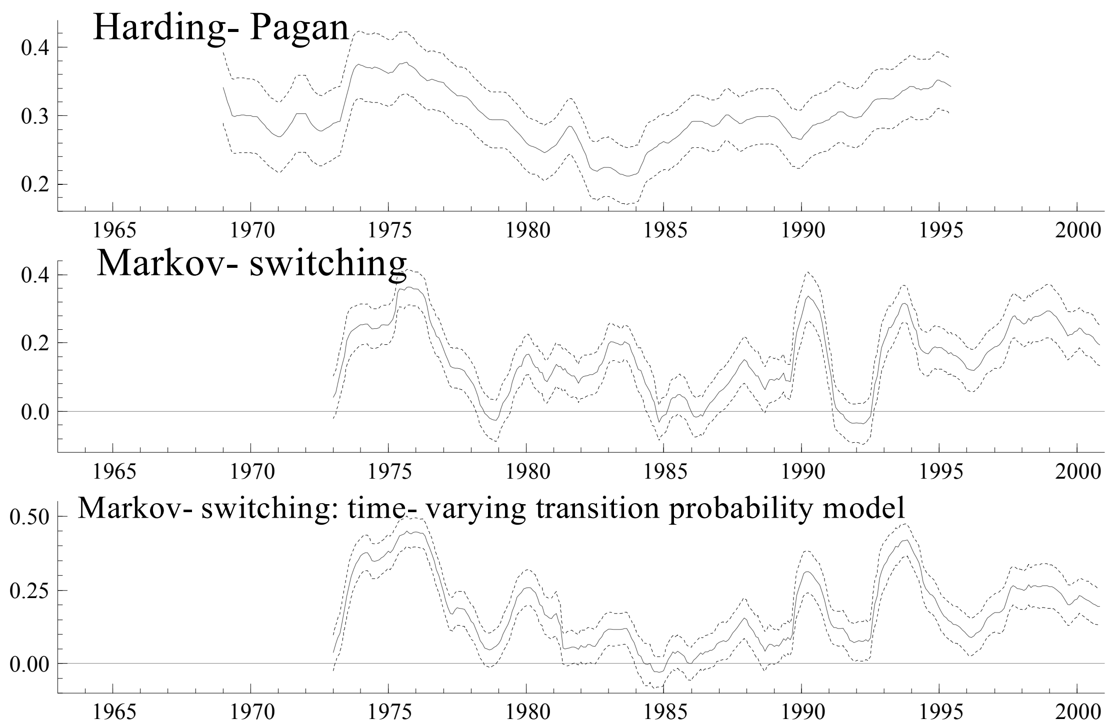

Figure 4: Average correlation between Euro-zone business cycles: three measures of the classical cycle

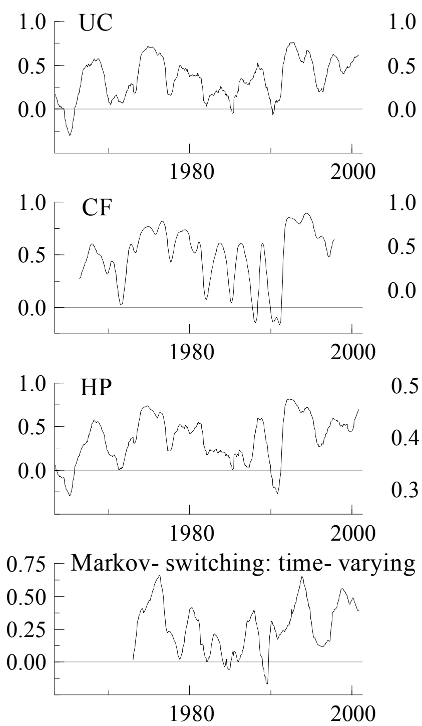

Figure 5: Average correlation between German and remaining Euro-zone business cycles for alternative measures of the business cycle

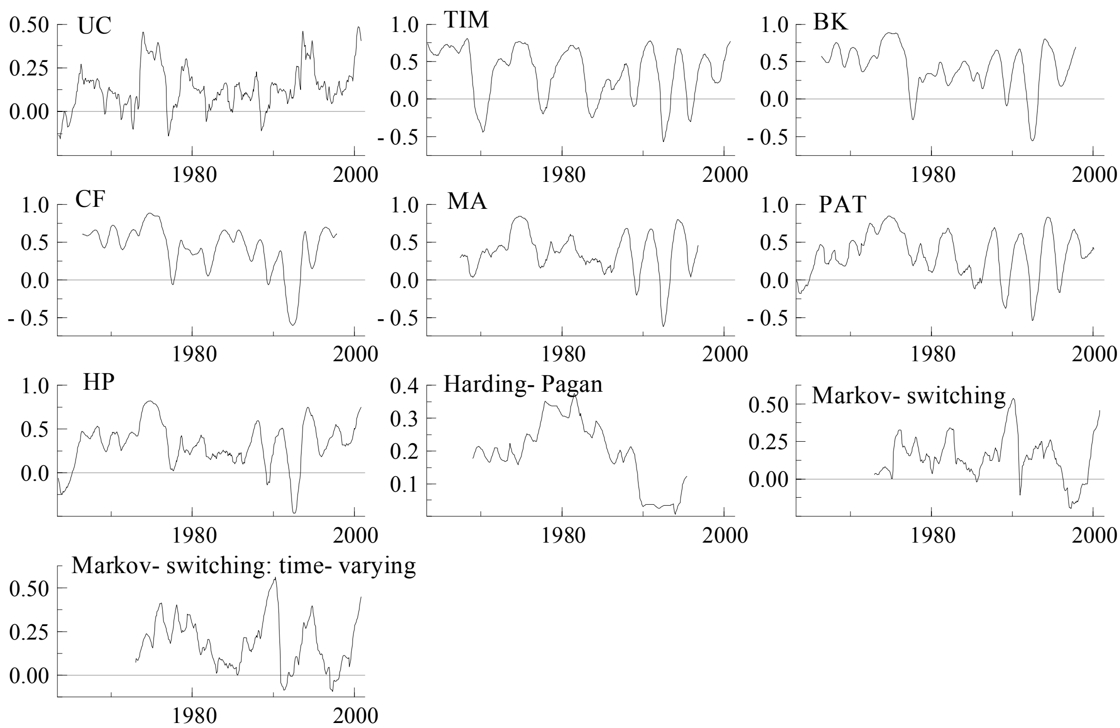

Figure 6: Average correlation between US and Euro-zone (excluding German) business cycles for alternative measures of the business cycle

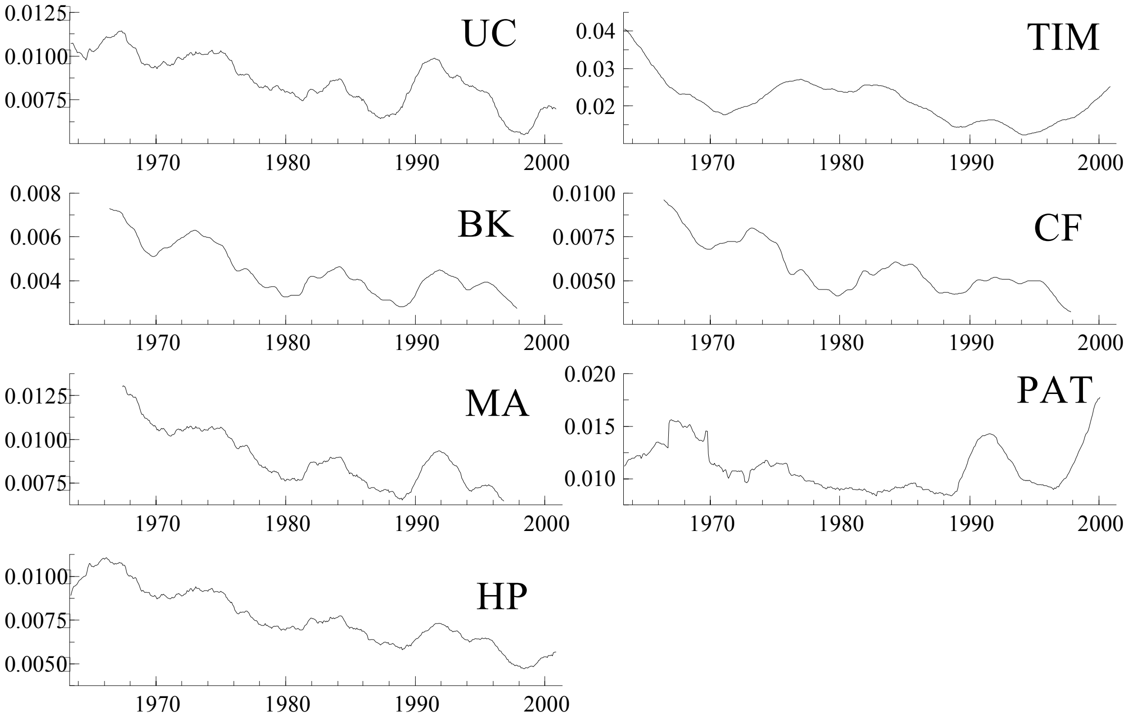

Figure 7: Average root mean squared difference between Euro-zone business cycles, for alternative measures of the growth cycle

EUROSTAT COLLOQUIUM
MODERN TOOLS FOR BUSINESS CYCLE ANALYSIS

Luxembourg, 28 and 29 November 2002
Session on "Multivariate Decomposition Methods"

THE COMMON CONVERGING TREND-CYCLE MODEL: ESTIMATION, MODELING AND AN APPLICATION TO EUROPEAN CONVERGENCE

BY ROB LUGINBUHL AND SIEM JAN KOOPMAN[1]

Department of Econometrics, Free University, NL-1081 HV Amsterdam

Tinbergen Institute, Amsterdam

This paper discusses multivariate time series models based on unobserved components with dynamic converging properties. We define convergence in terms of a decrease in dispersion over time and model this decrease via mechanisms that allow for gradual reductions in the ranks of covariance matrices associated with the disturbance vectors driving the unobserved components of the model. The inclusion of such convergence mechanisms within the formulation of unobserved components makes the identification of various types of convergence possible. For example, in a panel of macroeconomic time series for different countries, convergence in rates of growth, in cyclical behaviour and in overall volatility can be modelled separately and jointly. Each convergence mechanism introduces two, or in the case of overall volatility, three additional parameters. These parameters can be estimated simultaneously with the other parameters of the model. A mix of EM and numerical maximisation methods are used to obtain maximum likelihood estimates. The multivariate unobserved common converging component model is applied to the per capita gross domestic product for five European countries: Germany, France, Italy, Spain and the Netherlands.

KEYWORDS: Common trends and cycles; Dynamic factor model; Economic convergence; Kalman filter; State space; Multivariate unobserved components time series models.

JEL CLASSIFICATION: C13; C32; E32.

[1]Correspondence author: Department of Econometrics, Free University Amsterdam, De Boelelaan 1105, NL-1081 HV Amsterdam, The Netherlands. Tel: +31 20 444 6019 Fax: +31 20 444 6020 Email: `s.j.koopman@feweb.vu.nl`

1 Introduction

Unobserved components time series (UC) models typically consist of interpretable components such as trends, cycles, and seasonal and irregular components. Each component is separately modelled by an appropriate dynamic stochastic process which usually depends on normally distributed disturbances with mean zero and a given variance. The UC model with trend and cycle components enables a time series decomposition that is appropriate for many macroeconomic time series such as consumption, investment and national income. In a univariate time series analysis this model primarily leads to a particular trend-cycle decomposition and can be regarded as a model-based alternative to analyses based on the Hodrick-Prescott filter of Hodrick and Prescott (1980) and the Beveridge-Nelson decomposition of Beveridge and Nelson (1981). Harvey and Jaeger (1993) have argued that a model-based trend-cycle decomposition for economic time series is to be preferred and can avoid the detection of spurious cycles in the time series. Working with UC models has the additional advantage that they can also be used for producing forecasts.

The multivariate extension of the UC model can be used for simultaneous decompositions of a group of related time series. As a result the components become vectors defined by stochastic functions of vector disturbances generated by multivariate distributions. In the case of Gaussian models, disturbance densities rely on variance matrices rather than scalar variances. The application of non-diagonal variance matrices requires that the time series be modelled simultaneously. The design of these variance matrices is of interest from an economic standpoint. They determine, for example, whether the trends are positively correlated to each other, and the extent to which there are cycles common to a subset of the time series. From an econometric perspective, the multivariate extension of UC models is of interest because it enables the modeller to identify specific stable relationships between time series. It is therefore interesting to note that the phenomenon of cointegration can explicitly be modelled within an UC framework by including so-called common trends.

Several contributions in the economic literature can be found that study the phenomenon of convergence in the context of economic growth theory, see, for example, Williamson (1981), Galor (1981) and Quah (1981). Several definitions of convergence have been suggested and several inference procedures have been developed for the detection of convergence based on cross-section and time series data, see, for example, Bernard and Durlauf (1981) and the references therein. Many time series contributions use straightforward techniques for the comparison of two specific time series. In such analyses simple hypotheses are investigated such as whether the production series of, for example, the US and Japan have converged. A distinction is drawn between specific forms of convergence such as convergence of growth rates to the same level and convergence of overall variation.

We will adopt the definitions of convergence introduced by Barro and Sala-I-Martin (1992). Our starting point will be the definition of convergence as a reduction of the variation in a cross-section of time series. Whether the reduction of variability is due to harmonization in the underlying dynamics of the growth, the cyclical behaviour, the volatility or a combination of these is to be deducted from our modelling strategy. This paper does not focus on testing hypotheses of convergence such as is done in Bernard and Durlauf (1995). Another contribution is by Harvey and Carvalho (1980) who analyses convergence using unobserved components with balanced growth trends and concentrate on the formulation of convergence tests. In this paper we introduce multivariate time series models with explicit time-varying rank-reduction mechanisms will be formulated and estimated by maximum likelihood in order to identify convergence features present in the economic time series.

Convergence in this paper will be loosely defined and referred to as a reduction in dispersion over time. We propose to model convergence via the gradual reduction in the rank of the covariance matrix associated with the disturbance vector driving the appropriate unobserved component. The convergence mechanism that will be introduced has not been used within the multivariate UC framework and can therefore be regarded as a new development. The stochastic process governing a particular unobserved component can be made subject to the proposed convergence mechanism. This makes the identification of various types of convergence possible. For example, the short-term business cycle dynamics of a cyclical process may convergence, while the long-term dynamics of a trend may not. In general this approach permits the investigation of convergence to be directed towards the identification of which types of convergence are present. Therefore while it is possible that both the trend and cycle components may converge, it may equally well be the case that only a typical feature of the trend component, say, the growth of the trend, may be subject to convergence. In this paper we will present an illustration of a panel time series of the real gross domestic product (per capita, in logs) from five different European countries that appear to be subject to convergence in the rate of growth, cyclical behaviour and the overall variance. In our analysis we are able to identify these different types of convergence by defining a multivariate UC model with convergence mechanisms and by estimating the parameters using maximum likelihood methods.

The multivariate UC models with convergence can be represented as time-varying state space models with the bulk of the parameters arising in the variance structure of the model. The Kalman filter is the main tool to compute the loglikelihood for standard UC models with a given set of parameters. Maximum likelihood estimates of the parameters are obtained by initially using the EM method of estimation and continuing by directly maximising the log-likelihood with respect to the parameters using numerical Newton type optimization methods. The convergence UC model introduces two to three additional parameters for each convergence mechanism. The additional parameters can be estimated simultaneously with the other parameters. Specific adjustments are required for the EM method and for the evaluation of analytical scores.

The multivariate unobserved converging component model is applied to the per capita real gross domestic product for five European countries: Germany, France, Italy, Spain and the Netherlands. Various UC models both with and without convergence have been fitted and a full account of the modeling process is reported in this paper. We show that the main convergence features in these series is present in the cyclical pattern and the overall volatility.

The remaining part of the paper is organized as follows. In Section 2 we introduce the convergence mechanism for a simple multivariate UC model. A general time series model with trend and cycle components is discussed and is estimated for the European GDP series in Section 3. The converging common trend-cycle model is considered in Section 4. Empirical results for the European GDP series are presented in Section 5 and some final remarks are made in Section 6.

2 The converging local level model

We introduce the concept of unobserved converging components for the multivariate local level (LL) model which we regard as the simplest multivariare unobserved components time series model. The LL model is given by

$$y_t = \mu_t + \varepsilon_t, \qquad \mu_{t+1} = \mu_t + \eta_t, \qquad t = 1, \ldots, n, \tag{1}$$

where y_t is a $p \times 1$ vector of observed time series that is decomposed into the unobserved level component μ_t and the irregular vector ε_t. The level is modelled as a multivariate random walk process and the irregular is a vector white noise process. Further we assume Gaussian densities for the disturbance vectors

$$\varepsilon_t \sim \mathcal{N}(0, \Sigma_\varepsilon), \qquad \eta_t \sim \mathcal{N}(0, \Sigma_\eta), \qquad t = 1, \ldots, n,$$

and both disturbance vectors are serially and mutually uncorrelated. The time series characteristics and properties of model (1) are discussed in detail in Harvey (1989). The variance matrix Σ_η plays an important role in the interpretation of the LL model. When Σ_η is positive definite the model represents a set of seemingly unrelated time series equations (SUTSE). When the level variance matrix has a lower rank, the model can be represented by the common level model

$$y_t = a + A\mu_t^* + \varepsilon_t, \qquad \mu_{t+1}^* = \mu_t^* + \eta_t^*, \qquad t = 1, \ldots, n, \tag{2}$$

where a is a $p \times 1$ vector with the first r elements equal to zero, A is a $p \times r$ unity lower triangular matrix and level μ_t^* is a $r \times 1$ vector that represents a multivariate random walk process as in (1) with innovation vector $\eta_t^* \sim \mathcal{N}(0, D_\eta)$ and $r \times r$ diagonal matrix D_η. It can be shown that $\mu_t = a + A\mu_t^*$ and $\Sigma_\eta = AD_\eta A'$ with $\text{rank}(\Sigma_\eta) = r$. The variance matrices Σ_η and Σ_ε are unknown and need to be estimated by maximum likelihood. The evaluation of the likelihood function can be based on the output of the Kalman filter; see Durbin and Koopman (2001) for a recent overview of such state space methods. The likelihood function is numerically maximised with respect to the variance matrices. The matrices are typically decomposed so that $\Sigma_\eta = AD_\eta A'$. The elements in A and D_η are then estimated, thereby ensuring that the restriction that Σ_η is non-negative definite holds. The elements of D_η are estimated in logs to make sure that they are always non-negative.

System convergence in this paper is modelled by a gradual reduction in the rank of the variance matrices associated with the disturbances driving the unobserved components. For example, when the rank of Σ_η is r_1 at the beginning of the sample and $r_2 < r_1$ at the end of the sample, the local level model is said to be subject to system convergence in levels. This feature of convergence can be modelled in various ways. It can be introduced as a "structural" break within the rank of the variance matrix. A more realistic alternative, however, is to allow for a smooth transition towards a lower rank regime. In this paper we introduce a convergence mechanism that is parsimonious and elementary. Other specifications could also be considered within our framework.

The convergence mechanism for the ith element of D_η is specified by the deterministically time-varying logit function

$$d_{\eta,t,i} = \exp(2d_{\eta,i}^*) \exp(s_{\eta,t,i}) / \{1 + \exp(s_{\eta,t,i})\}, \tag{3}$$

where $d^*_{\eta,i}$ represents the logarithm of a standard deviation and therefore determines the size of the variance. The variable $s_{\eta,t,i}$ is given by

$$s_{\eta,t,i} = s^*_{\eta,i} \times (t + \tau_{\eta,i}), \tag{4}$$

where $s^*_{\eta,i}$ determines the rate of the variance change and $\tau_{\eta,i}$ determines the mid-timepoint of the change. According to the specification in (3), the variance matrix D_η, now $D_{\eta,t}$, is time-varying, as is the variance matrix Σ_η, which is replaced by

$$\Sigma_{\eta,t} = AD_{\eta,t}A', \qquad t = 1, \ldots, n.$$

In this specification the variance $d_{\eta,t,i}$ will converge to zero for any finite $d^*_{\eta,i}$, $s^*_{\eta,i}$ and $\tau_{\eta,i}$. Whether the convergence of $d_{\eta,t,i}$ to zero takes place within the sample range of $t = 1, \ldots, n$ depends on the values of $s^*_{\eta,i}$ and $\tau_{\eta,i}$. If a particular element of the time-varying variance matrix $D_{\eta,t}$ converges to (virtually) zero via the convergence mechanism for $t < n$, the rank of $\Sigma_{\eta,n}$ will be one less than the rank of $\Sigma_{\eta,1}$.

The convergence mechanism can be introduced for every diagonal element of $D_{\eta,t}$ and the estimation of the variance parameters, including the additional coefficients $s^*_{\eta,i}$ and $\tau_{\eta,i}$ for $i = 1, \ldots, r$, can take place simultaneously. However, the introduction of convergence mechanisms should take place after some prior analysis of the data. An illustration for an elaborated unobserved components time series model for a multiple time series of European gross domestic product (GDP) in logs is given in the next section.

3 Trend-cycle decompositions of European GDP series

Many economic time series typically feature a long term trend with cyclical variations around this trend. Further they are often characterised by trends with different growth rates for different periods and by cycles with time-varying characteristics. The identification of these unobservable features can be improved by considering a multiple set of similar time series. Therefore a multivariate unobserved components time series model is considered.

3.1 Trend and cycle components

An unobserved components model for a multiple set of non-seasonal economic time series can be expressed as

$$y_t = \mu_t + \psi_t + \varepsilon_t, \qquad t = 1, \ldots, n, \tag{5}$$

where y_t represents the actual time series, and the unobserved components consist of the trend μ_t, cycle ψ_t and irregular ε_t, all of which are vectors. These unobserved vectors are modelled as stochastic processes. For example, a simple model for the trend μ_t is given as in (1). However, the long-term trends of the GDP series are subject to positive growths which can be incorporated in the trend specification as follows,

$$\mu_{t+1} = \mu_t + \beta_t + \eta_t, \qquad \beta_{t+1} = \beta_t + \zeta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma^2\Sigma_\eta), \quad \zeta_t \sim \mathcal{N}(0, \sigma^2\Sigma_\zeta), \tag{6}$$

for $t = 1, \ldots, n$ and $\sigma^2 > 0$ where β_t is the $p \times 1$ vector of growth terms. This trend model is known as the local linear trend model. If $\Sigma_\eta = \Sigma_\zeta = 0$, then $\beta_{t+1} = \beta_t = \beta$, and $\mu_{t+1} = \mu_t + \beta$, so that the trend model (6) reduces to a deterministic linear trend. If $\Sigma_\eta = 0$ and $\Sigma_\zeta \neq 0$, the growth vector β_t is stochastically time-varying and the trend vector μ_t has become a cumulator function of growth terms which will result in a smooth trend component. Deficient ranks of variance matrices Σ_η and Σ_ζ imply common trend and growth components; see the discussion in Harvey (1989) and Harvey and Koopman (1997). The presence of common trends implies what is known as "cointegration" in the econometric literature; see Barro and Sala-I-Martin (1988). In short, it assumes that a $p - r_\eta \times p$ matrix of cointegrating vectors G exists such that Gy_t is stationary. This means that $GA = 0$.

Various specifications for a cycle component ψ_t may be considered. For example, we may wish to generate multiple cycles by vector autoregressive processes. Alternatively we can represent the cyclical processes by a set of trigonometric terms with time-varying coefficients. It can be shown that a stochastic cyclical process can be incorporated in a multivariate time series model. The multiple cycle component is given by

$$\begin{pmatrix} \psi_{t+1} \\ \psi_{t+1}^+ \end{pmatrix} = \rho \left(\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \otimes I_N \right) \begin{pmatrix} \psi_t \\ \psi_t^+ \end{pmatrix} + \begin{pmatrix} \kappa_t \\ \kappa_t^+ \end{pmatrix}, \tag{7}$$

where $c = \cos\lambda_c$, $s = \sin\lambda_c$ and I_k is the $k \times k$ identity matrix. The $p \times 1$ vector ψ_t consists of similar cycles that have a common frequency λ_c and a common autoregressive coefficient $|\phi| < 1$. The disturbance vectors are serially and mutually uncorrelated, and are normally distributed with mean zero and variance matrix

$$\mathrm{Var} \begin{pmatrix} \kappa_t \\ \kappa_t^+ \end{pmatrix} = I_2 \otimes \sigma^2 \Sigma_\kappa,$$

such that κ_t and κ_t^+ have a common variance matrix Σ_κ. This specification generates a stationary multiple cyclical process with a period of $f_c = 2\pi/\lambda_c$. The individual cycles in ψ_t have similar properties due to the common damping factor ρ and the cycle period λ_c. More details on similar cycles can be found in Harvey and Koopman (1997). The model is complete by taking the irregular component ε_t as a normally random vector with mean zero and variance matrix $\sigma^2\Sigma_\varepsilon$. The irregular and other disturbances associated with the various components are mutually uncorrelated, both contemporaneously and between different time periods. We finally note that $\sigma^2 > 0$ is the scaling variance constant or the common variance. It can be estimated when one of the non-zero elements in the variance matrices Σ_η, Σ_ζ, Σ_κ and Σ_ε is set to one. The common variance is introduced because we will allow it to vary over time in the next section in order to model variance convergence.

The unknown parameters in the trend plus cycle model are the variance matrices Σ_ε, Σ_η, Σ_ζ and Σ_κ together with the autoregressive coefficient ϕ and the cycle frequency λ_c. These parameters are estimated by maximum likelihood for which the Kalman filter is employed to compute the loglikelihood function for a given set of parameters. The fixed common variance can be concentrated out of the loglikelihood function and can be estimated implicitly by the Kalman filter output; see Harvey (1989). The resulting concentrated loglikelihood function can be maximised numerically with respect to the vector of parameters using particular Newton methods for nonlinear optimisation. The variance matrices need to be non-negative definite while the cycle coefficients are subject to the restriction $0 \leq \rho < 1$ and $\lambda_c = 2\pi/f_c$ with a cycle period of $f_c > 2$. The restrictions are implemented by representing the variance matrices via the Cholesky decompositions

$$\Sigma_\eta = AD_\eta A', \qquad \Sigma_\zeta = BD_\zeta B', \qquad \Sigma_\kappa = CD_\kappa C', \qquad \Sigma_\varepsilon = ED_\varepsilon E', \tag{8}$$

where the matrices A, B, C, E have unity lower triangular structures and their number of columns depends on the number of nonzero diagonal elements in the diagonal matrices D_η, D_ζ, D_κ, D_ε. To enforce the restriction that variance matrices must be positive definite we have

$$D_a = \text{diag}\left\{\exp(2d^*_{a,1}), \ldots \exp(2d^*_{a,r_a})\right\}, \tag{9}$$

for $a = \eta, \zeta, \kappa, \varepsilon$ and r_a is the rank of the corresponding variance matrix (assuming that the elements $d^*_{a,1}, \ldots, d^*_{a,r_a}$ do not tend to $-\infty$). Finally, we have $\rho = |\theta_\rho| \, (1 + \theta_\rho^2)^{-1/2}$ and $f_c = 2 + \exp\theta_\lambda$ where θ_ρ and θ_λ are the parameters that are actually estimated together with the matrices from the Cholesky decomposition of the variance matrices.

The Kalman filter is used for the computation of the loglikelihood function. This requires the model to be represented as a state space model. The state vector is given by

$$\alpha_t = (\mu_t', \beta_t', \psi_t', \psi_t^{*\prime})',$$

with initial conditions that are properly defined for the cycle vectors but are regarded as diffuse for the trend vectors μ_t and β_t since these elements represent nonstationary processes. The Kalman filter can account for diffuse initial conditions using extensions such as the ones discussed in Chapter 5 of Durbin and Koopman (2001). The computations are implemented using the `Ox` object-oriented matrix programming environment of Doornik (1999) and using the `Ox` library of `C` functions for state space models `SsfPack` by Koopman, Shephard, and Doornik (1993). The computations for the basic multivariate model of this section can also be carried out by the STAMP package of Koopman, Harvey, Doornik, and Shephard (1999).

3.2 Multivariate decomposition of European GDP series

We consider a multiple time series of real gross domestic production per capita (GDP) for five European countries: Germany, France, Italy, Spain and the Netherlands. The time series are measured in local currencies on a quarterly basis and covers the period of the first quarter of 1970 to the first quarter of 2001. We obtained the required population and GDP series from OECD sources[2] and calculated the series after which we standardized each series to the value of 100 in 1970. The rescaling of the data is justified due to the fact that the series are measured in different currencies. We model the logarithm of these time series multiplied by 100. The resulting time series is denoted by

$$y_t = \log \ \text{GDP} \begin{pmatrix} \text{Germany} \\ \text{France} \\ \text{Italy} \\ \text{Spain} \\ \text{Netherlands} \end{pmatrix} \text{ at time } t = 1, \ldots, n \text{ corresponding to } 1970Q1, \ldots 2001Q1,$$

with $n = 125$. The five time series are presented in Figure 1. The typical features of trends and cycles in the GDP series can be detected.

[2] See `http://www.sourceoecd.org/content/html/index.htm`.

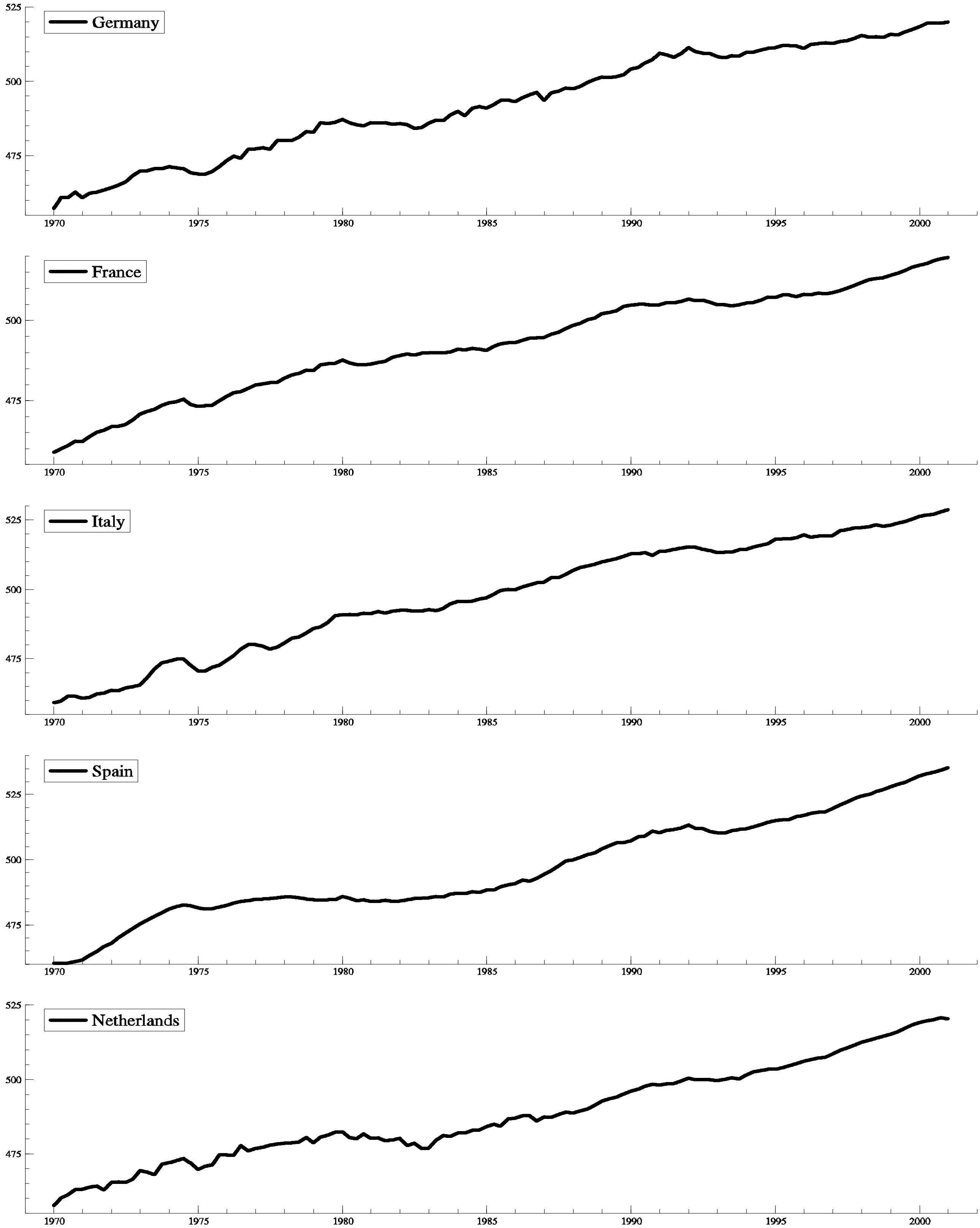

Figure 1: Quarterly seasonal adjusted real gross domestic production (per capita) for Germany, France, Spain, Italy and the Netherlands for the period 1970Q1 – 2001Q1

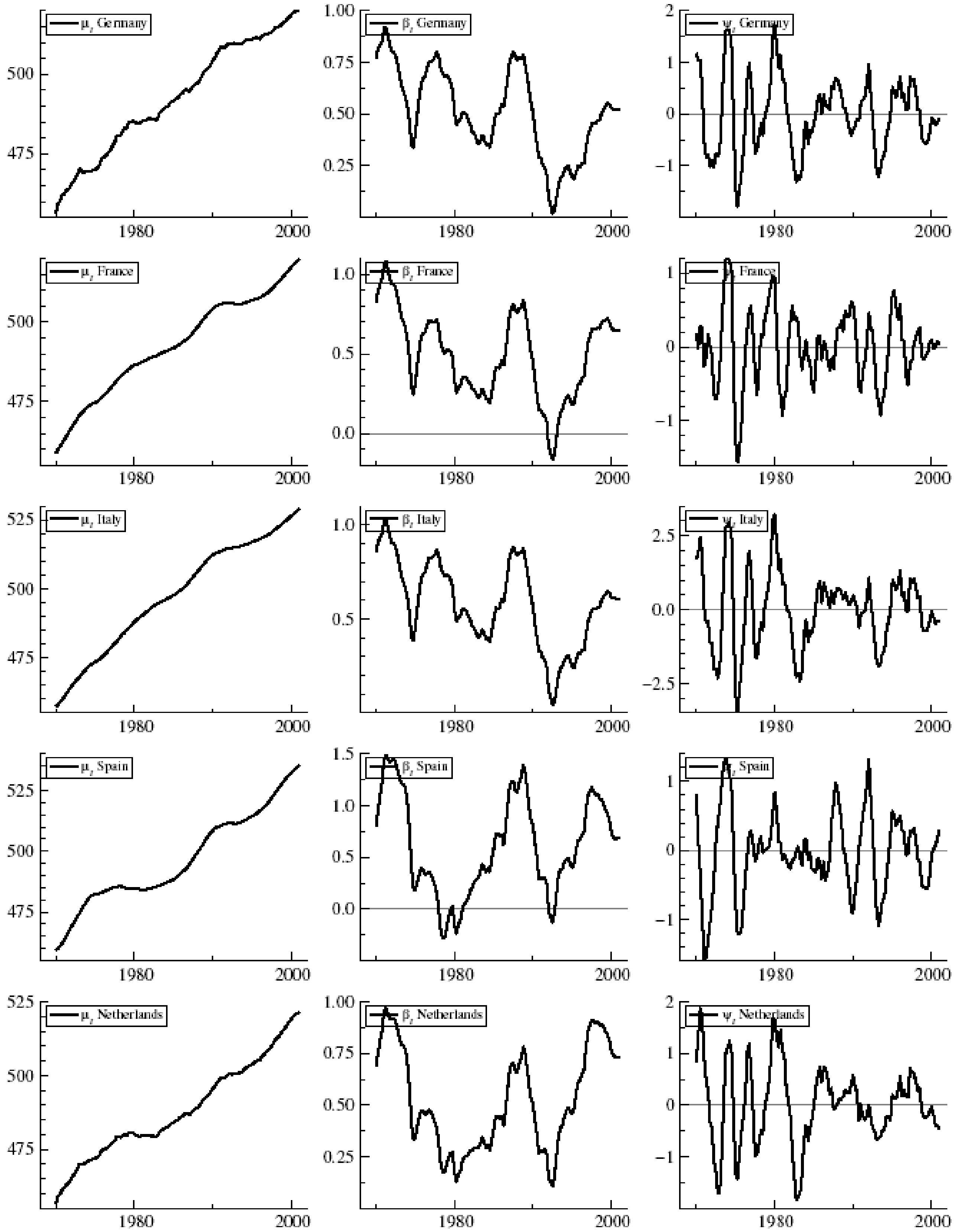

Figure 2: Trend, slope and cycle components for common trend-cycle model

The unobserved components model introduced in the previous section is estimated using the standard Kalman filter but modified to allow for diffuse initial conditions for the trend and slope components. The resulting loglikelihood is maximised numerically with respect to the parameters of the model which include the elements in the variance matrices and the autoregressive coefficient and the period of the similar cycle component. The estimated transformed parameters of the diagonal variance matrices D_ε, D_η, D_ζ and D_κ in (8) are reported in Table 1 in the third row of each panel (sample 1970–2001).

It follows from these results that the estimates of the trend variance matrices Σ_η and Σ_ζ have lower ranks. In particular, the estimates imply that $r(\Sigma_\eta) = 1$ and $r(\Sigma_\zeta) = 3$, indicating that the five trends are relatively smooth. This conclusion is confirmed by Figure 2 where the estimated trends are presented in the first column and the associated slope components are presented in the second column of graphs. Although there are some individual differences in the slope series, one can clearly see that the slopes for Germany, France and Italy are similar, and that this group of slopes for these three countries in turn is different from the similar pair of slopes for Spain and the Netherlands, although the patterns from 1986 onwards are similar for all five countries.

The estimated parameters for the multiple cycle component are given by

$$\hat{\theta}_\rho = 2.48 \quad \text{(s.e 0.38)}, \qquad \hat{\theta}_\lambda = 2.72 \quad \text{(s.e 0.09)},$$

which imply a similar cycle with an estimated autoregressive coefficient ρ of 0.93 (with a 95% confidence interval between 0.87 and 0.96) and an estimated cycle period f_c of 17.2 quarters (with a 95% confidence interval between 14.8 and 20.0).

The estimate of the variance matrix Σ_κ, which applies to both disturbance vectors κ_t and κ_t^+, has a rank of 3 according to the estimate of matrix D_κ reported in Table 1.

Table 1: Estimated diagonal variance matrices of common trend-cycle model

	sample	d_1^*	d_2^*	d_3^*	d_4^*	d_5^*
	1970-1986	-1.86	*	*	*	*
D_η	1987-2001	-1.36	*	-1.51	*	*
	1970-2001	-0.47	*	*	*	*
	1970-1986	-1.52	-2.41	-2.67	-2.11	*
D_ζ	1987-2001	-2.42	-2.59	*	*	*
	1970-2001	-2.30	-3.49	*	-2.44	*
	1970-1986	-1.17	-1.77	-2.34	*	*
D_κ	1987-2001	-1.47	-2.38	*	*	*
	1970-2001	-1.27	-1.51	-1.98	*	*
	1970-1986	-0.47	-1.87	-2.01	*	*
D_ε	1987-2001	-0.77	-1.80	-1.48	-1.94	*
	1970-2001	-0.70	-1.60	-1.74	*	-1.06

The estimates are for the transformed values of the diagonal variance matrices in (8) for three different samples. The transformed value d_i^* is the logarithm of the square root of the ith diagonal element of D. Large negative values produce therefore variances which are close to zero.

The cycles of the five countries can therefore be described by three mutually uncorrelated common similar cycles. These empirical results alone can be of interest to economists. European policy decisions on integration can benefit from the empirical confirmation of the existence of common cycles, especially if the actual dependence of individual European countries on common cyclical movements can be empirically identified.

To be able to adequately address this last point, we would need to report the estimate of the C matrix in (8). To limit the number of tables in this paper we will only present the estimates of the matrices A, B, C and E in the discussion of the empirical results for the converging models.

According to the diagnostic test statistics reported in Table 4 under the columns denoted by M, the estimated common components model is not entirely satisfactory. The table presents some standard diagnostics based on the residuals of the estimated model.

The standardized one-step ahead residuals are assumed to be standard normal distributed and serially uncorrelated. It follows from the reported statistics that France and the Netherlands may be subject to some moderate outliers, the dynamics are not well captured for Spain and for most series the residuals appear to be heteroskedastic. It is of interest to determine whether these model diagnostics improve when we consider converging models.

The common components model provides a satisfactory overall fit when we compare the sum of squared residuals with the sum of squared residuals from a model with only a constant and a fixed time trend. The percentage decrease of the common components model for all five series is at least 82% in relation to the naive model with a maximum decrease for Spain and a minimum decrease for Germany.

3.3 Preliminary evidence of convergence in European GDP series

We now focuses on the question of whether it is reasonable to assume that the rank of a variance matrix changes over time to a lower rank. If this is the case we will consider this to be an indication of some converging behaviour by a particular set of components. The full sample is split into the two roughly equivalent subsamples of 1970–1986 and 1987–2001. The trend-cycle components model is re-estimated for both of these subsamples. The estimation results of the variance matrices D, which determine the rank of the variance matrices in (8), are produced in Table 1. It is perhaps not surprising that we do indeed find evidence that convergence mechanisms play a role within the multiple time series. The estimated variance matrix of the slope component has a rank of 4 for the first sample, whereas in the second sample its rank is 2. Similarly, in the case of the cycle variance matrix in the first sample, the estimate of D_κ indicates a rank of 3 while in the last sample it indicates a rank of 2. These results provide some evidence of slope and cyclical convergence although any conclusions made must be tentative ones, because the subsamples consist of five series with at most 68 observations, representing only a moderate sample size. Furthermore, the cycles have an estimated period of more than 4 years so that realistically no more than three, or at most four full cycles can be observed within each subsample.

4 Unobserved common converging components

In this section we introduce the converging mechanism of Section 2 for the growth and cycle components. We therefore will represent the multivariate trend-cycle decomposition model as a dynamic factor model for the $p \times 1$ vector of time series y_t. We obtain

$$y_t = a + A\mu_t^* + C\psi_t^* + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2 \Sigma_\varepsilon), \tag{10}$$

where a is a $p \times 1$ vector with the first r_η elements equal to zero, A is a $p \times p$ unity lower triangular matrix and C is a $p \times r_\kappa$ unity lower triangular matrix. The matrix A differs from the A in (2) in that zero rows and zero columns, corresponding to the diagonal zero elements in D_η, must be added. All of the diagonal elements of A, however, including those of the added zero rows and columns, are unity. The specification of common growth rates given below in subsection 4.1 necessitates the augmentation of the matrix A in this manner. The matrices A and C are sometimes referred to as factor loading matrices. The unobserved factor μ_t^* represents the $p \times 1$ vector of underlying trends and factor ψ_t^* represents the $r_\kappa \times 1$ vector of cyclical components.

An unobserved component with an activated convergence mechanism (3) that models the gradual rank reduction over time is defined as an *unobserved converging component*. From an economic viewpoint, convergence in growth and cycle components can be of major interest as we will show in the analysis of European GDP series in Section 5 where models with unobserved converging components are estimated. In subsections 4.1 and 4.2 we introduce the convergence mechanisms for the trend and cycle component. In subsection 4.3 we turn our attention to the specification of variance convergence.

4.1 Common converging slope component

The trend component is specified as

$$\begin{aligned} \mu_{t+1}^* &= \mu_t^* + b + B^*\beta_t^* + \eta_t^*, & \eta_t^* &\sim \mathcal{N}(0, \sigma^2 D_\eta), \\ \beta_{t+1}^* &= \beta_t^* + \zeta_t^*, & \zeta_t^* &\sim \mathcal{N}(0, \sigma^2 D_\zeta), \end{aligned} \tag{11}$$

where the $r_\zeta \times 1$ vector β_t^* consists of the underlying growth terms of the trend μ_t^* for $t = 1, \ldots, n$. The initial values for μ_1^* and β_1^* are treated as being generated from diffuse density functions. The vector b is a $p \times 1$ vector of unknown constants of which the first r_ζ are fixed at zero. The $p \times r_\zeta$ factor loading matrix B^* is unity lower triangular. The disturbance vectors η_t^* and ζ_t^* are mutually and serially uncorrelated and have diagonal variance matrices D_η and D_ζ, respectively. Note that the matrix B in 8 is given by $B = AB^*$, and therefore the matrix Σ_ζ can be expressed as $\Sigma_\zeta = AB^* D_\zeta B^{*\prime} A'$.

When $r_\eta = r_\zeta = p$ and $\psi_t = 0$ in (10), the multivariate local linear trend model reduces to a SUTSE specification in which the variance matrices Σ_η and Σ_ζ are of full rank. Common trend specifications are obtained when $r_\eta < p$ or $r_\zeta < p$ or both. It is noted in Section 2 that common trends imply cointegration.

In the application of this paper, the convergence of the trend takes place in the growth rate. We will therefore give the details of convergence for the slope component. A gradual change of a non-zero value for a particular element of D_ζ to a zero value can be accomplished by the logit function

$$d_{\zeta,t,i} = \exp(2d^*_{\zeta,i}) \exp(s_{\zeta,t,i})/\{1 + \exp(s_{\zeta,t,i})\}, \qquad i = 1, \dots, r_\zeta, \tag{12}$$

where $d_{\zeta,t,i}$ is the ith element of the time-varying variance matrix $D_{\zeta,t}$ that replaces D_ζ in (11). The size of the logarithm of the standard deviation of the ith element of $D_{\zeta,t}$ is $d^*_{\zeta,i}$ for $i = 1, \dots, r_\zeta$. The variable $s_{\zeta,t,i}$ is specified as

$$s_{\zeta,t,i} = s^*_{\zeta,i} \times (t + \tau_{\zeta,i}), \qquad i = 1, \dots, r_\zeta,$$

where $s^*_{\zeta,i}$ determines how quickly the function $s_{\zeta,t,i}$ approaches zero and $\tau_{\zeta,i}$ determines the mid-point of this change over time. The same converging mechanism can be introduced for the variance matrix of the trend component itself, $D_{\eta,t}$, which is then also time-varying and replaces D_η in (11). The details of this convergence mechanism are the same as for the slope component and are discussed in Section 2.

4.2 Common converging cycle component

The cycle component ψ^*_t is defined in a similar way as in equation (7) with $\psi_t = C\psi^*_t$ and $\psi^+_t = C\psi^{*+}_t$ where

$$\begin{pmatrix} \psi^*_{t+1} \\ \psi^{*+}_{t+1} \end{pmatrix} = \rho \left(\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \otimes I_{r_\kappa} \right) \begin{pmatrix} \psi^*_t \\ \psi^{*+}_t \end{pmatrix} + \kappa^\times_t, \qquad \kappa^\times_t \sim \mathcal{N}(0, I_2 \otimes \sigma^2 D_\kappa), \tag{13}$$

for $t = 1, \dots, n$. The cycle vectors ψ^*_t and ψ^{*+}_t have dimension $r_\kappa \times 1$ and the vector of disturbances $\kappa^\times_t$ has dimension $2r_\kappa \times 1$. The fact that $\psi_t = C\psi^*_t$ such that $\Sigma_\kappa = CD_\kappa C'$ is only valid for similar cycles; see Harvey and Koopman (1997) for the technical details.

The convergence mechanism introduced for the trend component in the previous subsection can also be incorporated in the diagonal variance matrix of the cycle disturbances D_κ. It implies that matrix D_κ becomes time-varying and its i-th element will be modelled in the same manner defined in (12) for $i = 1, \dots, r_\kappa$.

4.3 Common converging variance component

The specification of a common converging variance component as a formalization of convergence in the overall variance requires three parameters. Unlike the other forms of convergence for the trend, slope, and cycle components, the common variance σ^2 cannot be allowed to converge to zero unless we are willing to entertain the idea of convergence to a deterministic system of non-stochastic equations. To ensure that we retain some volatility after the convergence of the common variance, we introduce a third constant parameter into the convergence specification. This leads to the following specification for variance convergence,

$$\sigma^2_t = \exp(2d_\sigma) + \exp(2d^*_\sigma) \exp(s_{\sigma,t})/\{1 + \exp(s_{\sigma,t})\}. \tag{14}$$

According to (14), the total volatility before convergence sets in is given by $\exp(2d_\sigma)+\exp(2d^*_\sigma)$, while only the constant term $\exp(2d_\sigma)$ remains after convergence has occurred. Both d_σ and

d_σ^* are the logarithm of a standard deviation. The size of the constant component of the total volatility is determined by d_σ, while d_σ^* determines by how much volatility declines due to variance convergence. In an analogous fashion to the definition of $s_{\eta,t,i}$ in (4), the variable $s_{\sigma,t}$ is specified as

$$s_{\sigma,t} = s_\sigma^* \times (t + \tau_\sigma)$$

where s_σ^* determines the rate at which the function $s_{\sigma,t}$ approaches zero and τ_σ determines the mid-point of the change over time.

5 Convergence in European GDP series

We now turn our attention to the question of how the various types of convergence presented in the previous section can be applied in practice to develop an appropriate statistical model. We demonstrate this by generalizing the trend-cycle model presented in Section 3 to include the three different types of convergence.

5.1 Specification and estimation of converging trend-cycle model

To apply the convergence mechanisms, we first determine which diagonal elements $d_{a,i}$ from the matrices D_a for $i = 1, \ldots, r_a$ and $a = \eta, \zeta, \kappa$, are candidates to converge to zero. Given the structure of the Cholesky decompositions of the variance matrices in (8), it is possible to infer a number of simple rules that will hold in most situations. We illustrate these rules via a discussion of D_ζ.

Generally, it will not be the case that the first diagonal element $d_{\zeta,1}$ will converge to zero. Although technically possible, the result of such convergence is that the variance of the disturbance driving the slope component for the first series degenerates to zero. In other words, the element $d_{\zeta,1}$ represents the sole source of variability for the slope component of the first series. This can be verified directly via (8). In the case of our model, $d_{\zeta,1}$ corresponds to the variance of the slope component for Germany. Examination of the estimated values given in Table **??** for $d_{\zeta,1}$ suggests that this parameter is not subject to the convergence mechanism, given that its value is the largest one in the D_ζ matrices for each sample. The same can be said of the values for $d_{\eta,1}$ and $d_{\kappa,1}$.

The second diagonal element $d_{\zeta,2}$ represents the marginal contribution to the variance of the slope disturbance for the second series, France, after accounting for the contribution made by $d_{\zeta,1}$. The contribution of $d_{\zeta,1}$ is the result of the correlation between the slope disturbances for the first, Germany, and second series, France.

In general, the structure of the Cholesky decomposition implies that the variance of the slope disturbance for series j is only affected by $d_{\zeta,1}, \ldots, d_{\zeta,j}$. The elements $d_{\zeta,i}$ for $i > j$ make no contribution to the variances of the slope disturbances for the series j. Therefore, in the case of the second series, France, it is also less likely that $d_{\zeta,2}$ is subject to the convergence mechanism, because this would imply that the slope disturbances of Germany and France become perfectly negatively or positively correlated after convergence has taken place.

We wish to point out that it is a relatively strong statement about slope convergence for these two series. This is, of course, not to suggest that this could not have happened, but it is unlikely. In any case, the results in Table 1 indicate that the estimated value of $d_{\zeta,2}$ remains roughly constant over the entire sample period. This also applies to the estimates shown for $d_{\kappa,2}$, while in the case of $d_{\eta,2}$ all estimates indicate that France does not have its own variance for the trend disturbance in any of the sample periods.

For the series i for $i > 2$, the convergence of $d_{\zeta,i}$ to zero implies that the variance of the slope disturbance for the series i becomes a linear combination of slope disturbances of the first $i-1$ series. For this reason the correlations between the slope disturbance for the series i with those for the first $i-1$ series will still typically be less than one in absolute value after convergence.

We therefore regard the last element d_{ζ,r_ζ} as the most likely candidate for convergence. This is a consequence of the fact that the variance of the slope disturbance for the series r_ζ becomes a linear combination of all other values of $d_{\zeta,i}$ in the case that this element converges. Generally it is the case that the larger the number of elements that are involved in a linear combination, the greater the chance one element will converge. We therefore conclude for $a = \zeta, \kappa$ that the most likely candidate for convergence is $d_{a,i}$ for the largest available value of i, that is $i = r_a$. Inspection of the estimates in Table 1 supports this conclusion as well.

Where there is evidence of $d_{a,i}$ dropping out of the model in the latter part of the sample period for $a = \eta, \zeta, \kappa$, it is for values of i of either 3 or 4. In the case of $d_{\eta,i}$, the estimates show no evidence of convergence. In fact the estimates obtained for the entire sample period indicate that only one value $d_{\eta,1}$ is responsible for the trend volatility for all five countries in Σ_η. For this reason we do not apply the convergence mechanism to D_η and restrict our attention to model specifications in which $r_\eta = 1$. This has the advantage of resulting in more parsimonious models, which, given the number of model parameters, is an important advantage.

Returning to the example of the slope component, we note that a further consequence of the Cholesky decomposition of Σ_ζ is that the variances and correlations among the slope disturbances for the first $i-1$ series are not influenced by the existence of a convergence mechanism on $d_{\zeta,i}$. The variances of the slope disturbances for the series j, where $j \geq i$, can be altered by this convergence.

The correlations between the slope disturbances for the series j, where $j \geq i$, with the slope disturbances from all other series can also be changed by convergence. For this reason we also consider it to be important, in the case of our example of GDP, to place the largest, most dominating and stable economies first in the observation vector, with the smaller, more dependent, and less stable economies last, because the latter economies are more likely to have been effected by convergence towards the economies of the more dominating countries of the European Union (EU). Although this rule does not lead to a unique ordering of the countries in the observation vector y_t, it does in any case seems reasonable to have Germany and France, two of the original member of the EU, as the first two elements of y_t, given the economic dominance of these countries. Italy too is one of the original members of the EU and also has a larger economy than either Spain or the Netherlands, although smaller than that of either Germany or France. For these reasons we opted to place Italy third in the observation vector. We chose to place Spain fourth in the vector, reflecting the fact that Spain only joined the EU in 1986. Although the Netherlands has been a member of the EU since its inception, it is a smaller country with little power to influence the larger economies in the EU, and as a result, the Netherlands is the last series.

In summary, based on the estimates for the common trend-cycle model without convergence mechanisms reported in Table 1, as well as on the model structure implied by the combination of the Cholesky decomposition parameterization of the variance matrices together with the convergence mechanism, we have opted to produce estimates for the trend-cycle model of the previous section with the incorporation of converging mechanisms for the slope β_t, the cycle ψ_t and the common variance σ_t components. At first the convergence mechanism is introduced to element 4 of the diagonal variance matrix of the common slope component β_t and the resulting model is indicated by M_β. In the same way, the convergence mechanism is placed on the third element of the diagonal matrix D_κ to investigate cycle convergence only. This model is indicated by M_ψ. The trend-cycle model without any convergence mechanism is indicated by M and the model with only the convergence mechanism for σ^2 is indicated by M_σ. Although we have also investigated employing the convergence mechanism on other elements of D_ζ and D_κ, no other specification produced significant results based on the loglikelihood values and standard information criteria such as the Akaike information criterion (AIC). In the discussion of the results below we only consider the models with convergence mechanisms that are estimated significantly. Finally, the model with convergence considered for all components will be indicated by $M_{\beta\psi\sigma}$.

The full modelling and estimation process is as follows. First the non-converging model M is estimated for which some of the results for the European GDP series are reported and discussed in Section 3. The estimated parameters of the M model act as the starting values for the parameters of the models M_β, M_ψ and M_σ. In particular, we maintain the restrictions of $r_\eta = 1$ and $r_\varepsilon = 4$ in all model specifications in the interests of parsimony. Initial values for τ and s^* of the convergence mechanisms are chosen so that the original values of D_ζ, D_κ or σ remain almost fixed throughout. We then explore the possibility that convergence can take place at various points in time. This amounts to setting s^* close to zero. We typically started with a value of $\tau = n/2$, but we also thoroughly explored values corresponding to the start date for the European Exchange Rate Mechanism in 1979Q1 and Spanish entry to the EU in 1986Q1. The initial values used in the maximum likelihood optimization routine for the converging models are similar to the values for M, some of which are reported in Table 1. The estimates for the models M_β, M_ψ and M_σ are then used to initialize the maximum likelihood routines for the models employing two type of convergence: $M_{\beta\psi}$, $M_{\beta\sigma}$ and $M_{\psi\sigma}$. Finally, estimates obtained for the the latter more complex converging models serve in turn as starting values for the final and most complex model of interest $M_{\beta\psi\sigma}$.

All models, including the converging trend-cycle models, are estimated by maximum likelihood for which the Kalman filter is used for the evaluation of the loglikelihood function. The Kalman filter can handle time-varying state space models which we require for representing the converging model in state space form since the variance matrices are time-varying due to the convergence mechanisms defined in (12) and (14).

5.2 Estimation results for European GDP series

The estimation results for model M are discussed in Section 3 and the estimated elements of the diagonal variance matrices of the common components are reported in Table 1. The parameters in Table 1 correspond to the d^* parameters of the converging mechanism defined in (12) and (14) and their estimates for the three single convergence models are reported in Table 2 together with the convergence rate parameter s^* and the timing parameter τ. Estimates of the same parameters for the final model $M_{\beta\psi\sigma}$ are also reported. The models M_σ and $M_{\beta\psi\sigma}$ incorporate variance convergence which requires a third convergence parameter d_σ. This parameter is also reported in Table 2.

Since the variance matrix Σ_η for the level disturbance η_t is estimated to have rank one in model M, we have used this restriction for all converging models. The estimated value of $d^*_{\eta,1}$ is around -0.5 for most models except for M_σ. In fact, the estimated d^* values in D_ζ and D_ε for the various convergence models do not vary a great deal. This indicates that a large number of parameters in the convergence models are estimated in a numerically stable manner.

Of particular interest are the estimates of the parameters s^* and τ of the converging mechanisms. The actual values of s^* and τ are not straightforward to interpret and therefore we graphically present in Figure 3 the three different convergence mechanisms for the slope, cycle and common variance components which are estimated simultaneously for model $M_{\beta\psi\sigma}$.

The graphs clearly show two different converging patterns. By the end of the sample period the slope and cycle convergence processes have taken place, while the common variance is still in the process of converging. It is estimated that the standard deviation of the variance will eventually converge to the value of 0.037. We note, however, that the final converged value of this variance is difficult to estimate accurately given that both the starting and converged values lie well outside the sample period. The midpoints of convergence for the slope is approximately 1981Q2 and for the cycle it is approximately 1983Q3, or about two years later. This fact will be discussed further below.

The log-likelihood values of the various estimated models, denoted by $\log(L)$, are reported in Table 3 together with the number of parameters, n_p, that are present in the model and the Akaike information criterion (AIC) that we have computed as

$$\text{AIC} = -2\log(L) + 2n_p. \tag{15}$$

We use the minimum of the AIC to determine the "best" model within the class of converging trend-cycle models in terms of fit relative to the number of parameters required for the estimation.

The minimum value of the AIC is found for the model $M_{\psi\sigma}$. This may indicate that cycle and common variance convergence is most relevant for the European GDP series. Although the slope convergence is strong, it is only relevant for the GDP of Spain and it has a lesser impact on the multiple GDP time series as a whole. However, for the illustrative purposes of this paper we will mainly discuss the estimation results of the full converging model $M_{\beta\psi\sigma}$ because it considers all aspects of the converging trend-cycle model.

The factor loading matrices A, B, C and E of the model $M_{\beta\psi\sigma}$ are estimated as

$$\widehat{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.08 & 1 & 0 & 0 & 0 \\ -0.42^* & 0 & 1 & 0 & 0 \\ -0.02 & 0 & 0 & 1 & 0 \\ 0.51^* & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \widehat{B} = \begin{pmatrix} 1 & 0 & 0 \\ 1.48^{**} & 1 & 0 \\ 1.03^* & 0.14 & 0 \\ 1.91^{**} & 1.39^{***} & 1 \\ 0.94^* & 1.33^* & 0.27^* \end{pmatrix}$$

$$\widehat{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0.48^{*} & 1 & 0 \\ 1.20^{**} & 0.49 & 1 \\ 0.77^{**} & -0.17 & -0.28^{*} \\ 0.32^{*} & -0.31 & 1.52^{*} \end{pmatrix}, \qquad \widehat{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0.18^{*} & 1 & 0 \\ 0.05 & -0.27 & 1 \\ -0.06 & -0.16 & -1.04^{***} \\ -0.24 & 1.37^{**} & -0.46^{*} \end{pmatrix}$$

where *, ** and *** indicate that the t-statistic is estimated as being greater than 1, 2 and 3, respectively. Zero values in the loading matrix A are the result of the fact that the corresponding values of d_i are 0. This causes these parameters to drop out of the model, hence the reported values of 0. There are some restrictions which could be applied by assigning various parameters a value of 0, particularly in the matrix E. This will not be explored further.

It can be inferred from the factor loadings of the converging element of the slope component (that is the third column of B) that the slope convergence is strongest for Spain.

The mid-point of the convergence is in the year 1982. Around this year the import and export activity of Spain with Europe increases exponentially while imports and exports with the countries of South America, for which traditionally strong trade links existed, hardly increases in this period.

For example, European Union (EU) imports were five times higher in 1992 than it was in 1980 (for exports it was four times). This has heavily influenced the development of GDP in Spain and its dependence on the EU member countries' economies. Therefore it may not be surprising that Spanish growth converges during this period toward the growth pattern followed by the EU member countries. This mainly explains the fact that the slope component converges from three factors to two factors.

There is also some weak evidence of the Dutch growth rate converging. This may partly be explained by the fact that Dutch GDP heavily depends on trade figures and more trade within Europe therefore likely leads to a stronger dependence on EU GDP growth. Finally it is noted that the fact that the factor loading matrices are restricted to be lower triangular and the fact that the fourth common slope component is converging, rather than the third, is evidence that the GDP growth of Germany, France and Italy converge before the beginning of the sample, that is before 1970.

The GDP of these three countries can be described by two common slope factors for the whole sample. The GDP growth of Spain and the Netherlands roughly converges to these two factors by 1986, the year in which Spain officially joined the EU.

The characteristics of the cyclical convergence are more complex, because more countries are involved. In fact we can distinguish two main GDP cycles in mainland Europe: one for Germany and one for France. The GDP cycles of the other three countries considered in our study converge towards these two cycles with full convergence taking place by the beginning of the 1990s.

Based on the estimated C matrix we can conclude that the five countries have the first common cycle component in common. The second common cycle is only significant for France since the factor loadings for the second cycle are not significant for Italy, Spain, or the Netherlands.

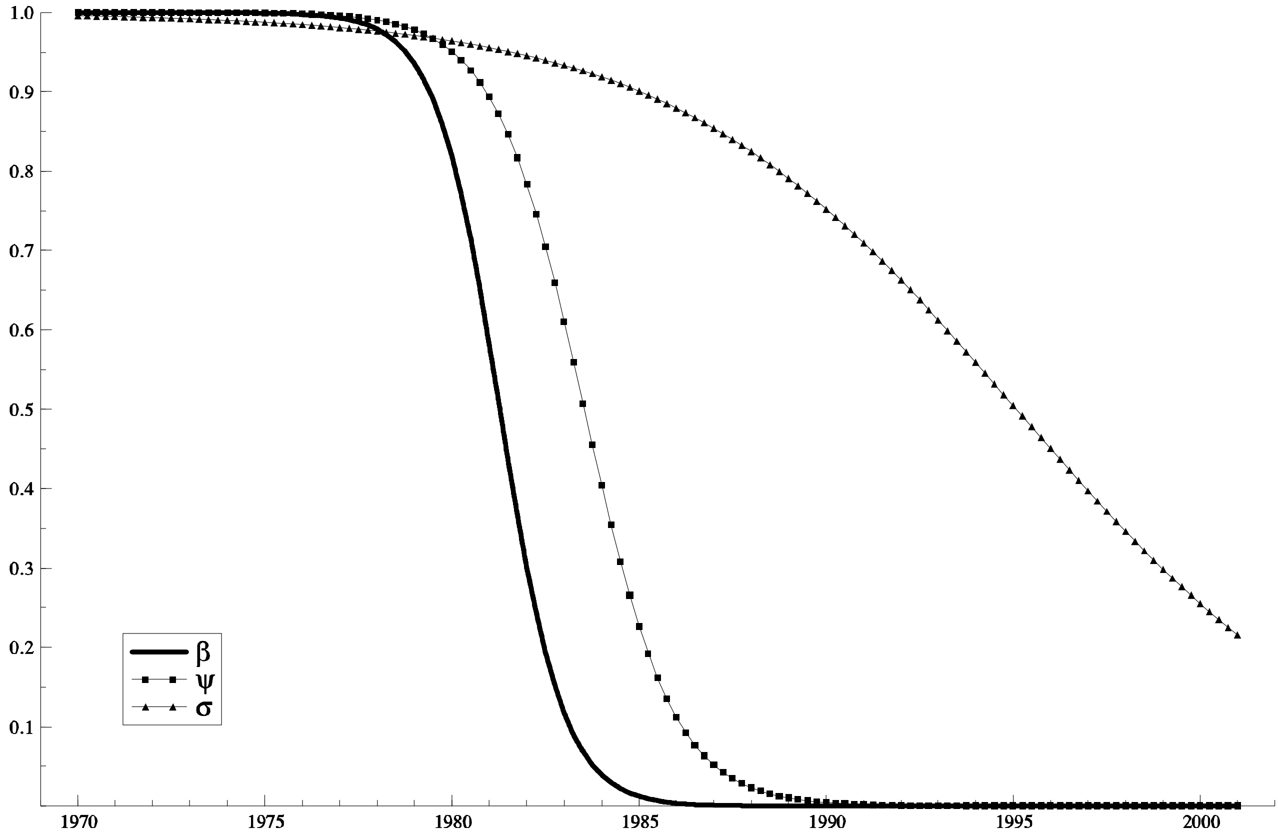

Figure 3: The converging mechanisms for slope (β), cycle (ψ) and variance (σ) convergence in model $M_{\beta\psi\sigma}$

Further, the converging cycle (the third column of C) is significant for Italy and the Netherlands, and to a lesser extent, for Spain. The empirical findings of cyclical convergence for Italy may well correspond to the inception of the Exchange Rate Mechanism. Arguably, stable economic conditions were required for entering the Exchange Rate Mechanism in 1979, as well for the later participation in the Euro, and as a result Italy was forced to follow the two dominating EU business cycles more closely. Given the discipline required by the goal of monetary union, it is even surprising that France deviates from the economic European cycle. However, the variation of the specific cycle for France (after correction for the European cycle) is moderate given its logged standard deviation estimated as -1.39 compared to the one of "Europe", that is -0.97, for model $M_{\beta\psi\sigma}$.

Finally, the estimated Cholesky matrix E of the irregular variance matrix in (8) is not of any significant interest. However, it is worth mentioning that the irregular series of Italy and Spain are strongly negatively correlated implied by the highly significant value of -1.04 in E, element $(3, 4)$. This may well be explained by the competitive nature of the trade of both countries within Europe. When Spain is doing well in terms of GDP, Italy may have done less well as a result, and vice-versa. Casual observation would certainly seem to suggest that various products and services traded in Europe come from both Italy and Spain, for example, wine and tourism.

Table 2: Estimated diagonal variance matrices of converging trend-cycle model

UC	model	d_1^*	d_2^*	d_3^*	d_4^*	d_5^*	i	s_i^*	τ_i	time
D_η	M_β	-0.53	*	*	*	*	-	-	-	-
	M_ψ	-0.50	*	*	*	*	-	-	-	-
	M_σ	-0.33	*	*	*	*	-	-	-	-
	$M_{\beta\psi\sigma}$	-0.45	*	*	*	*	-	-	-	-
D_ζ	M_β	-2.30	-3.66	*	-1.47	*	4	-0.43	48.2	1982Q1
	M_ψ	-2.18	-2.80	*	-2.34	*	-	-	-	-
	M_σ	-2.18	-2.92	*	-2.33	*	-	-	-	-
	$M_{\beta\psi\sigma}$	-2.30	-2.93	*	-1.45	*	4	-0.29	45.2	1981Q2
D_κ	M_β	-1.17	-1.48	-1.87	*	*	-	-	-	-
	M_ψ	-1.19	-1.46	-3.22	*	*	3	-0.50	66.2	1986Q3
	M_σ	-0.99	-1.42	-3.23	*	*	-	-	-	-
	$M_{\beta\psi\sigma}$	-0.97	-1.39	-0.86	*	*	3	-0.21	54.1	1983Q3
D_ε	M_β	-0.69	-1.61	-1.74	*	-1.01	-	-	-	-
	M_ψ	-0.65	-1.63	-1.95	*	*	-	-	-	-
	M_σ	-0.60	-1.41	-1.52	*	-0.99	-	-	-	-
	$M_{\beta\psi\sigma}$	-0.54	-1.49	-1.34	*	-0.87	-	-	-	-
σ_t^2	M_σ	d_σ=-3.16		$d_{\sigma,t}^*$=-0.60			-	-0.055	93.3	1993Q2
	$M_{\beta\psi\sigma}$	d_σ=-3.29		$d_{\sigma,t}^*$=-0.54			-	-0.055	100.4	1995Q1

The estimates are for the transformed values of the diagonal variance matrices in (8) for four different models of convergence. See Table 2 for the definition of the four models. The transformed value d_i is the logarithm of the square root of the ith diagonal element of D. Large negative values produce therefore variances which are close to zero. To facilitate the comparison of the results for the various models, the values for d_i obtained for models with variance convergence are from the first quarter of 1970 and are reported after re-scaling the values to reflect the contribution due to σ^2. This amounts to the rescaling of the variance so that $\sigma_1^2 = 1$.

Table 3: Log-likelihood values of estimated models

model	description	$\log(L)$	n_p	AIC
M :	UC, 1970-2001	-634.69	43	1355.37
M_{70-86}:	UC, 1970-1986	-371.22	45	-
M_{87-01}:	UC, 1987-2001	-184.80	42	-
M_β:	UCC, slope convergence	-628.42	45	1346.85
M_ψ:	UCC, cycle convergence	-616.72	45	1323.45
M_σ:	UCC, variance convergence	-604.28	46	1300.57
$M_{\beta\psi}$:	UCC, slope and cycle convergence	-609.62	47	1315.24
$M_{\beta\sigma}$:	UCC, slope and variance convergence	-602.08	48	1300.15
$M_{\psi\sigma}$:	UCC, cycle and variance convergence	-595.78	48	1287.57
$M_{\beta\psi\sigma}$:	UCC, slope, cycle, and variance convergence	-595.31	50	1290.63

The log-likelihood values are computed via the Kalman filter that is adapted for the diffuse initialisations of the trend and slope (nonstationary) components.

5.3 Diagnostic checking

The diagnostics of the standardised residuals (obtained for the estimated converging trend-cycle model $M_{\beta\psi\sigma}$ for normality, heteroskedasticity and serial correlation are presented in Table 4. They are for the most part satisfactory for the five residual series. In fact when they are compared with the diagnostics of the estimated non-converging model M they have generally improved.

In particular, the heteroskedasticity tests and the normality tests are all satisfactory except for Spain. The residual series of Spain indicate that a more modest decline in the variance may be required for the observations after 1986.

Figure 4 presents the cumulative sum of squared residuals for all standardised prediction residuals, that is

$$C_j = c_j/c_n, \qquad c_j = \sum_{t=1}^{j} v_t' F_t^{-1} v_t,$$

for $j = 1, \dots, n$, where v_t is the vector of one-step ahead prediction errors and F_t is its variance matrix at time t. Both v_t and F_t are computed by the Kalman filter.

Relative large deviations of C_j from j, whether they are positive or negative, indicate some structural break in the variance. It is clearly seen from Figure 4 that for the non-converging model M such breaks occur right from the beginning of our sample. This is the main reason for including the common variance converging mechanism of (14) in the final model $M_{\beta\psi\sigma}$.

The graph of C_j for the estimated model $M_{\beta\psi\sigma}$ in Figure 4 shows that the inclusion of variance convergence in the model has been effective in dealing with the heteroscedasticity in the time series due to breaks and other irregularities.

Table 4: Diagnostic checking for residuals of common converging trend-cycle model

	N_{DH}		$Q(10,7)$		$H(40)$		DW	
	M	$M_{\beta\psi\sigma}$	M	$M_{\beta\psi\sigma}$	M	$M_{\beta\psi\sigma}$	M	$M_{\beta\psi\sigma}$
Germany	1.24	0.34	7.58	8.72	0.28^{**}	0.56^{*}	1.86	1.89
France	7.85^{*}	2.79	11.7	14.1	0.33^{**}	0.72	1.79	2.00
Italy	0.37	1.63	11.2	5.09	0.32^{**}	0.96	1.48^{**}	1.62^{*}
Spain	3.03	5.66	23.9^{**}	14.3^{*}	0.88	3.03^{**}	1.91	1.86
Netherlands	6.25^{*}	3.97	8.22	9.82	0.24^{**}	0.89	1.93	1.75

M is the common trend-cycle model without convergence and $M_{\beta\psi\sigma}$ is the common converging trend-cycle model for the slope, cycle and variance components. The statistics are computed for the five standardized one-step ahead residuals obtained from the Kalman filter. N_{DH} is the asymptotic χ^2_2 normality test of Doornik and Hansen (1994). $Q(p,q)$ is the Box-Ljung test for the p autocorrelation and is asymptotically χ^2_q distributed. $H(k)$ is the standard heteroskedasticity test computed as the ratio of the sum of the first k and the sum of the last k squared residuals and is asymptotically $F(k,k)$ distributed. The notation * indicates significant at the 5% level and ** indicates significant at the 1% level.

5.4 Multivariate decomposition into trends and cycles

Given that the model diagnostics are satisfactory for the $M_{\beta\psi\sigma}$ model, we now present the estimated trend, slope and cycle components based on all observations (smoothed estimates of elements of the state vector). In Figure 2 we present the estimated vector components for the common trend-cycle model M and in Figure 5 the same smoothed estimates of the components are presented for the converging model $M_{\beta\psi\sigma}$. Comparison of the two figures indicates that they are similar as may be expected. In the case of the similar cycle component, the estimated parameters for the $M_{\beta\psi\sigma}$ model are also close to those reported for the M model in Section 3. For the $M_{\beta\psi\sigma}$ model, we obtained the following maximum likelihood estimates,

$$\hat{\theta}_\rho = 2.09 \quad \text{(s.e 0.29)}, \qquad \hat{\theta}_\lambda = 2.75 \quad \text{(s.e 0.10)}.$$

These values imply an estimated autoregressive coefficient ρ of 0.90 (with a 95% confidence interval between 0.83 and 0.94) and an estimated cycle period f_c of 17.7 quarters (with a 95% confidence interval between 14.9 and 21.1). There are, however, some subtle differences in the estimated components obtained with the two models.

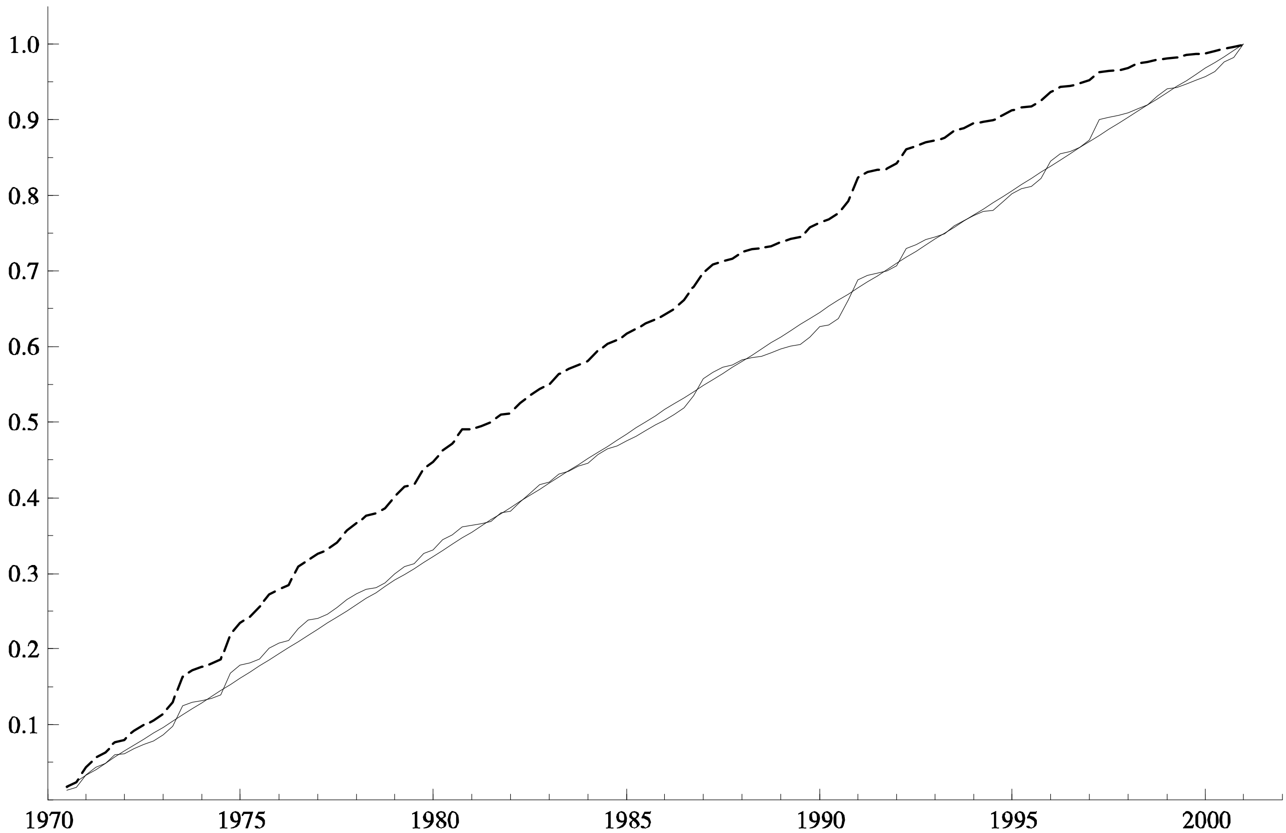

Figure 4: Cumulative sum of squared residuals C_j for common trend-cycle (M, broken line) and converging trend-cycle model ($M_{\beta\psi\sigma}$, solid line) together with the diagonal reference line (dotted line)

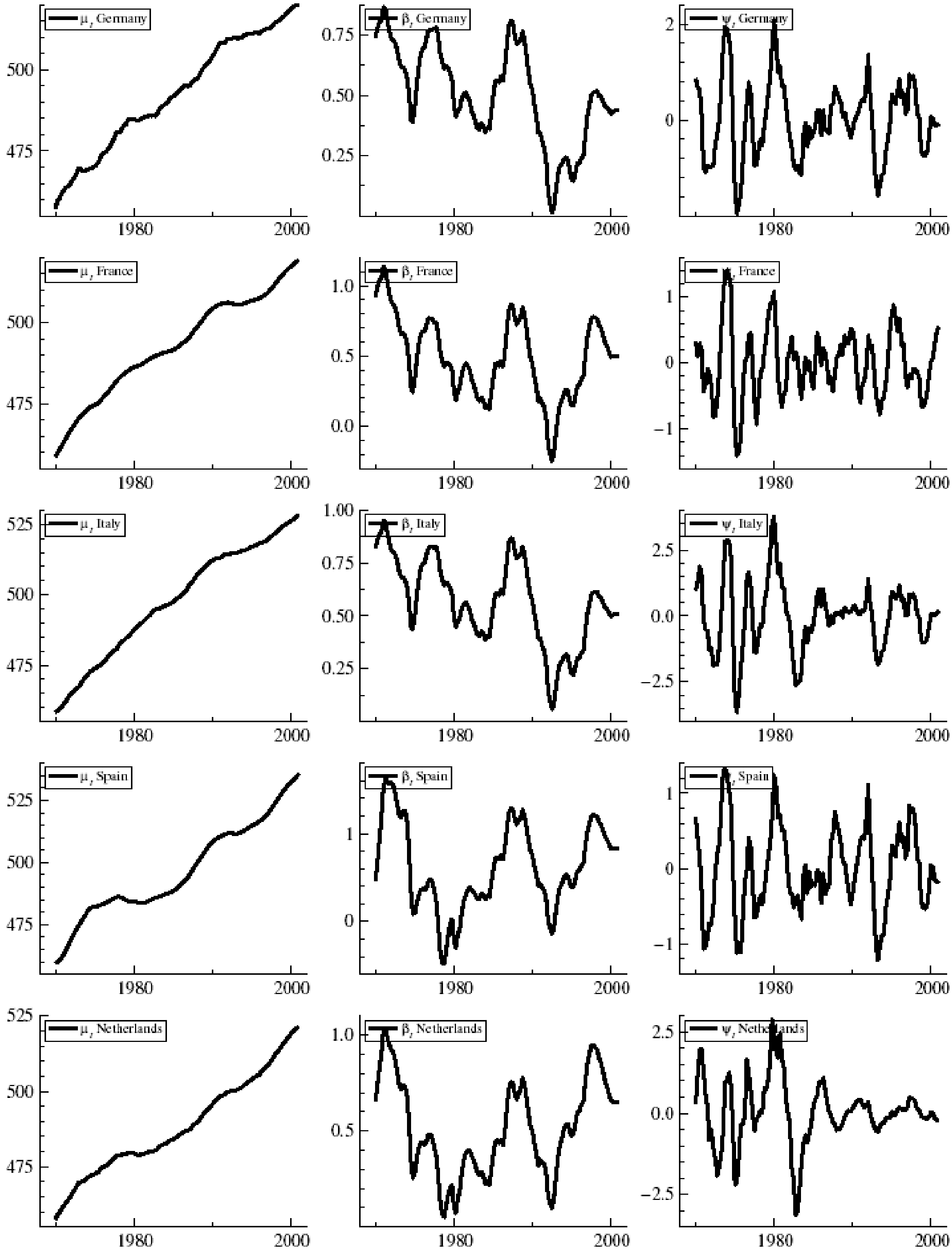

Figure 5: Trend, slope and cycle components for common trend-cycle model

The estimates of the slope and cycle components in the last ten years of the sample look more similar for the model $M_{\beta\psi\sigma}$ than those from the model M. This is to be expected because the slope and cycle components are both linear functions of only two underlying factors for the model $M_{\beta\psi\sigma}$ whereas for the model M they remain functions of three factors.

5.5 Time-varying correlations for converging common components

A quantity of particular interest for converging trend-cycle models is the correlation between the individual elements of the converging component. Let us define the time-varying level correlations by

$$\varrho_{\eta,t,i,j} = \sigma_t \Sigma_{\eta,t}(i,j) / \sqrt{\{\sigma_t \Sigma_{\eta,t}(i,i) \sigma_t \Sigma_{\eta,t}(j,j)\}}, \qquad (16)$$

where $\Sigma_{\eta,t} = AD_{\eta,t}A'$, the diagonal variance matrix $D_{\eta,t}$ being time-varying due to the convergence mechanism introduced in (3). Further, $\Sigma_{\eta,t}(i,j)$ refers to the (i,j) element of variance matrix $\Sigma_{\eta,t}$. These time-varying correlations can be presented as graphs and they provide information about which two countries are converging to each other in terms of level dynamics.

Similar quantities to those defined in (16) can be introduced for the slope ($\varrho_{\zeta,t,i,j}$) and cycle ($\varrho_{\kappa,t,i,j}$) components. Note that the correlations are not affected by variance convergence since the common variance σ_t cancels out in the computation of $\varrho_{\eta,t,i,j}$ and in the other correlations.

In Figure 6 we present a selection of the time-varying correlations for the slope and cycle components that in our view are of most interest. The slope correlations are presented for Germany, France, Italy, and Spain against Spain and the Netherlands. As pointed out earlier, the slope convergence is mainly due to Spain as the graphs make clear. The correlation in the slope disturbance between Spain and Germany is 0.62 at the beginning of the 1970s and it becomes 0.94 at the end of the 1980s which is a remarkable increase in about 15 years. Similar increases can be observed for the correlations of the slope disturbances with France and Italy. Slope correlation increases are also found for the Netherlands but on a smaller scale.

The correlation increases in the cyclical component are due to various countries as we have mentioned earlier. It is interesting to see that the cyclical correlations between Germany and the Netherlands increases in the 1980s from 0.2 to 0.8, between France and Italy from 0.57 to 0.79 and between Italy and Spain from 0.4 to 0.9. The most dramatic increase, however, is due to the change in the correlation between Spain and the Netherlands: from −0.18 to 0.91.

A closer economic investigation may be required to provide some explanation for these particular changes in the cycle correlations. However, the figures clearly show that a closer integration of the dynamics of the GDP cycle can be identified beginning around the time of the start of the European Exchange Rate Mechanism in March of 1979, and in the case of the slope convergence, ending around January 1986 when Spain became a member of the EU. The cycle component converges more slowly, but still before the opening of the EU's Common Market in January of 1993. That the cycle component converges more slowly is also evidenced by the fact that the parameter governing the rate of convergence for the cycle, $s^*_{\kappa,3} = -0.21$ is closer to 0 than the rate for the slope convergence, $s^*_{\zeta,4} = -0.29$, as can be seen in Table 2. Finally, however, we wish to sound a note of caution about the exact interpretation of these rates of convergence.

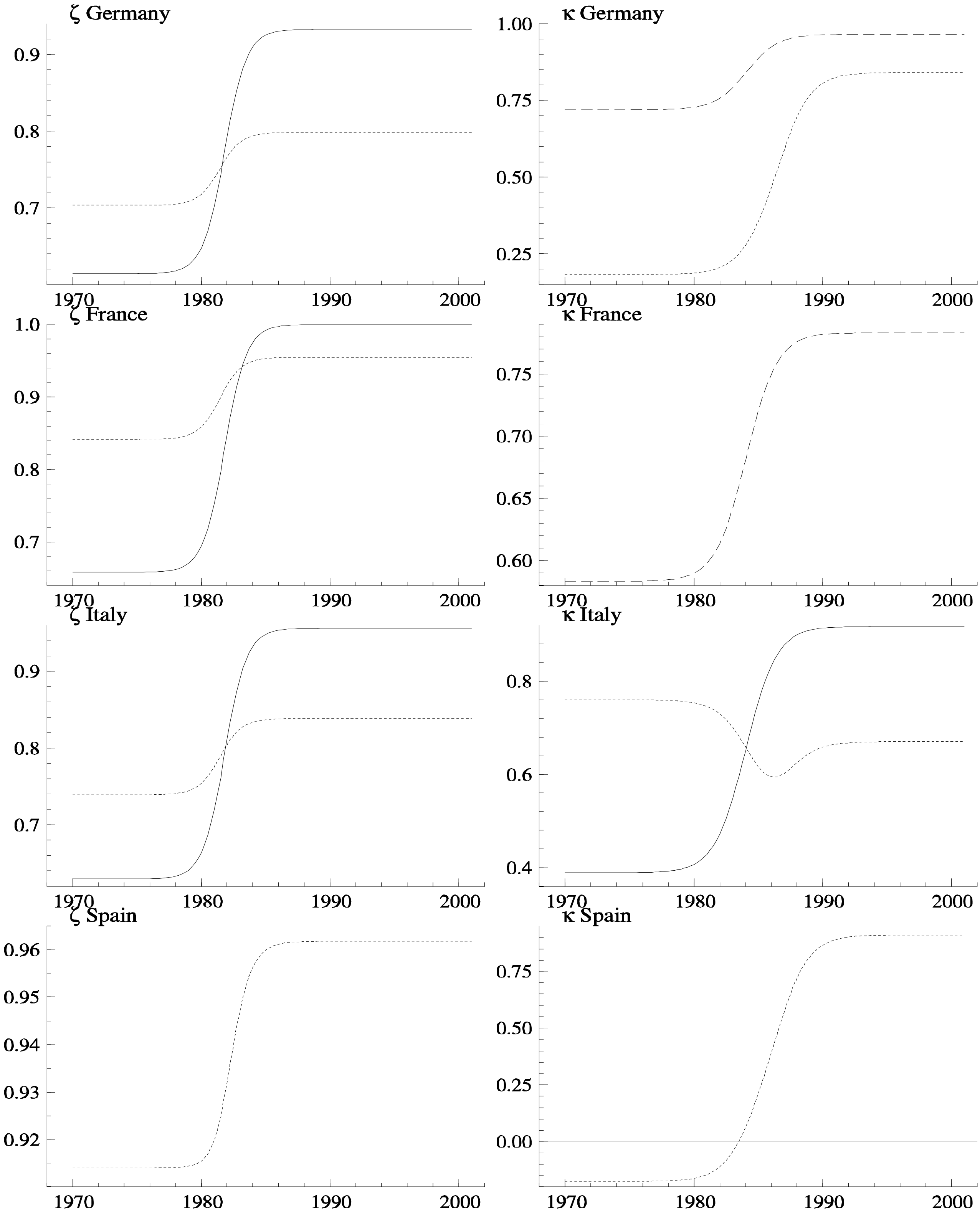

Figure 6: Some time-varying correlations $\varrho_{\zeta,t,i,j}$ and $\varrho_{\kappa,t,i,j}$ implied by the slope variance matrix $\Sigma_{\zeta,t}$ and the cycle variance matrix $\Sigma_{\kappa,t}$, respectively, for Spain (solid line), Netherlands (dotted line) and Italy (dashed line)

Our experience with estimating differing model specifications and differing sample periods has demonstrated that the estimates of these rates can vary as can been seen in Table 2 where the reported rate $s^*_{\kappa,3}$ for the model M_κ is further from 0 than the one for $s^*_{\zeta,4} = -0.29$ estimated for the model M_ζ. Nonetheless, both the slope and cycle components are consistently estimated as converging in the period around the early to mid 1980's.

6 Conclusions

In this paper we have considered a multivariate unobserved components model for quarterly GDP series of five European countries from 1970 to 2001. The model decomposes the $p \times 1$ observation vector, for $p = 5$, into unobserved trend, cycle and irregular components. We refer to this model as the trend-cycle decomposition model. Each component depends linearly on $r \leq p$ independent common factors. Multiple time series with $r < p$ common trends imply cointegrating relationships within the set of time series. Common cycles can be introduced into the model in a similar fashion. This leads to a model framework that can incorporate stable relationships between economic variables for the long term (trends) and the middle term (business cycles or other stationary components).

The main contribution of this paper is the introduction of convergence mechanisms into the common trend-cycle model. At the beginning of the time series, for example, the vector cycle component is a linear function of three factors, and subsequently converges to being dependent on only two factors. The process of rank-reduction is modelled as a smooth logistic function of time that depends on a shape parameter and a parameter that determines the mid-point of the convergence process. Such a convergence mechanism is introduced not only for the slope and cycle components, but also for the overall variance of the model in an adjusted form. The multivariate converging trend-cycle model is estimated by maximum likelihood via the Kalman filter and the results are discussed in detail.

On the basis of the empirical results, we draw the following three main conclusions.

1. Both the slope and cycle components begin converging following the introduction of the Exchange Rate Mechanism introduced by EU countries in 1979.

2. Spanish GDP growth converges to the common growth components of the other European countries by 1986 when Spain joined the EU.

3. The cyclical variations of the GDP series for Italy, Netherlands and Spain converge to the cycle processes of Germany and France by the beginning of the 1990's, before the inception of the Common Market in 1993.

The convergence of the Spanish growth rate can be explained by the huge increases of Spanish imports from and exports to other EU countries during the 1980s as Spanish prepared for entry to the EU, which occurred on 1 January 1986. The cycle convergence can be explained by the introduction of the Exchange Rate Mechanism in 1979 which demanded strong monetary coordination between participating EU countries, as well as by the necessary macroeconomic adjustments required by the EU member countries for the successful implementation of the Common Market in 1993, and the introduction of the Euro which followed. The convergence of the cycle at the beginning of the 1990's can therefore be taken as a consequence of the economic preparations for entry to the Common Market in the EU. Finally, we also included a convergence mechanism for the overall variance of the trend-cycle model. Although the estimated convergence of the common variance is significant, it had not fully converged by the end of the sample period in 2001, and according to our estimates is still in the process of converging.

References

Barro, R. J., and X. Sala-I-Martin (1988): "Testing for common trends," *Journal of the American Statistical Association*, 83, 1097–1107.

——— (1992): "Convergence," *Journal of Political Economy*, 100, 223–251.

Bernard, A., and S. Durlauf (1981): "Interpreting tests of the convergence hypothesis," *Journal of Econometrics*, 71, 161–173.

——— (1995): "Convergence in international output," *Journal of Applied Econometrics*, 10, 97–108.

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, 151–174.

Doornik, J. A. (1999): *Object-Oriented Matrix Programming using Ox*. Timberlake Consultants Ltd, London, 3rd edn., See http://www.nuff.ox.ac.uk/Users/Doornik.

Durbin, J., and S. J. Koopman (2001): *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford.

Galor, O. (1981): "Convergence ? Inferences from theoretical models," *The Economic Journal*, 106, 1056–1069.

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

Harvey, A. C., and V. Carvalho (1980): "Models for converging economies," Discussion paper, Faculty of Economics, University of Cambridge.

Harvey, A. C., and A. Jaeger (1993): "Detrending, stylised facts and the business cycle," *Journal of Applied Econometrics*, 8, 231–247.

Harvey, A. C., and S. J. Koopman (1997): "Multivariate structural time series models," in *Systematic dynamics in economic and financial models*, ed. by C. Heij, H. Schumacher, B. Hanzon, and C. Praagman. John Wiley and Sons, Cambridge.

Hodrick, R. J., and E. C. Prescott (1980): "Postwar U.S. business cycles: an empirical investigation," Discussion paper 451, Carnegie-Mellon University.

Koopman, S. J., A. C. Harvey, J. A. Doornik, and N. Shephard (1999): *Stamp 6.0: Structural Time Series Analyser, Modeller and Predictor*. Timberlake Consultants Ltd, London.

Koopman, S. J., N. Shephard, and J. A. Doornik (1993): "Statistical algorithms for models in state space using SsfPack 2.2," *Econometrics Journal*, 2, 113–166.

Quah, D. T. (1981): "Twin peaks: growth and convergence in models of distribution dynamics," *The Economic Journal*, 106, 1045–1055.

Williamson, J. G. (1981): "Globalization, convergence and history," *Journal of Economic History*, 56, 277–306.

Eurostat Colloquium
Modern Tools For Business Cycle Analysis

Luxembourg, 28 and 29 November 2002
Session on "Multivariate Decomposition Methods"

STATE SPACE DECOMPOSITION UNDER THE HYPOTHESIS OF NON ZERO CORRELATION BETWEEN TREND AND CYCLE, WITH AN APPLICATION TO THE EURO-ZONE

By Tommaso Proietti[1]

Dipartimento di Scienze Statistiche, University of Udine and Department of Economics, European University Institute

This paper discusses several issues related to trend-cycle decompositions with correlated components of macroeconomic time series, and illustrates them with reference to the Euro area and the Italian gross domestic product. In particular, we address the small sample properties of the estimated correlation of the trend and cycle disturbances, and review the interpretative issues raised by these decomposition. The nature of inferences about trends and cycles, with reference to the real time and final estimates, and the related topic of revision, is considered, along with the relationship with other popular results, such as the Beveridge and Nelson decomposition, the Single Source of Error and the Innovation models. We also look at the consequences of seasonal adjustment and temporal aggregation on the empirical evidence for a negative correlation between the disturbances. Finally, we illustrate that multivariate analysis can provide additional insight on this topic.

Keywords: Deviation and growth Rate Cycles; Hysteresis; Kalman filter and smoother; Innovation models; Seasonal adjustment; Temporal aggregation.

[1] Address for correspondence: Dept. of Economics, European University Institute - Badia Fiesolana, Via dei Roccettini 9, San Domenico di Fiesole (FI) I-50016 Tel. +39 055 4685 759. E-mail: tproiett@iue.it

1 Introduction

THIS PAPER IS CONCERNED WITH unobserved components (UC) models for the decomposition of a macroeconomic aggregate into a trend component and a *deviation* cycle. Unobserved components (UC) models assume that the components are driven by orthogonal disturbances (see, for instance, Clark, 1987, Harvey and Jäger, 1994) or perfectly correlated ones; in the latter case there is a single source of disturbances and the sign of the correlation is implied by the remaining parameter estimates. These restrictions are often enforced to produce just-identified decompositions, but in some cases they are over-identifying.

When there are no degrees of freedom available for estimating the correlation between the disturbances, and economic theory is uninformative about this parameter, we cannot usually discriminate between different assumptions by the usual likelihood inferences, although departures from the maintained model will show up in the several diagnostic tools in the econometrician kit. The reduced form, i.e. the corresponding model in the ARIMA class, will be of the same order, but models with orthogonal disturbances impose severe restrictions on its parameter space.

To overcome the latter, models with correlated components have been considered by Godolphin (1976) and Godolphin and Stone (1980), with the explicit intent of extending the parameter range yielding decomposable models. Snyder (1985), Ord, Koehler and Snyder (1997), and Hyndman *et al.* (2002) advocate state space models with only one source of random disturbances, with the same intent, also arguing that inferences are simplified.

Another very popular result, the Beveridge and Nelson (BN, 1981) decomposition, is formulated in terms of perfectly correlated disturbances, and is commonly viewed as providing a structural interpretation to *any* ARIMA model. In the study of macroeconomic time series (e.g. GDP at constant prices) the common disturbance has often been associated with productivity, or real, shocks.

The BN decomposition is actually a particular case of what is known as a *formal* decomposition of an ARIMA model; see Brewer *et al.* (1975), Brewer (1979) and Piccolo (1982). Casals, Jerez and Sotoca (2002) have recently advocated the use of the innovation form of a structural model for inference about unobserved components. Their argument is that this representation yields an “exact" decomposition, such that the model for the estimated component is congruent with the theoretical one and the components are estimated in real time, i.e. using only current and past information.

Models with perfectly correlated disturbances pay a price for their wider applicability: at the outset there is no guarantee that the components will be sensible. For instance, we must be willing to accept that trend growth has higher unconditional variance than output growth, which may be regarded as quite implausible, under a weak smoothness prior. Moreover, if main motivation for entertaining them was to enlarge the reduced form decomposable parameter range, but they happened to select a point in that space for which an orthogonal decomposition is admissible, the estimation of the components could have been improved by using future observations; smoothing is however prevented by the model specification itself.

In other words, the restrictions imposed by UC with uncorrelated components are often reasonable, providing plausible ways of weighting the data and of avoiding, for instance, that the trend fluctuates more wildly than the observations. Harvey and Koopman (2000), looking at the implications on the weighting patterns for signal extraction, cast some doubts on the plausibility of models with correlated components.

These issues have been reflected in the forecasting literature. The “structural" interpretation of the forecast function of any ARIMA model has been provided by Box, Pierce and Newbold (1987), using a partial fraction expansion of the autoregressive polynomial; by this approach, which is essentially the same as that at the foundation of the BN decomposition, it is possible to derive updating equations for the components of the forecast function that depend on the innovations. The fundamental question is whether the components can be genuinely interpreted as trends, cycles, etc., especially when the innovations are not discounted; see Proietti (2002a, sec. 5.10) for an illustration. On the other hand, UC models with orthogonal disturbances impose some kind of discounting on the innovations that is coincident with or in the same spirit of that arising in exponential smoothing techniques. With reference to the latter, the related problem as to whether smoothing constant greater than 1 are admissible has received some attention, and is reviewed in Gardner (1985).

A different situation arises when the restrictions on the correlation are over-identifying. This occurs if the representation chosen for the components is not “saturated": usually, UC models are a linear combinations of individual components, such that the i-th component has a (possibly nonstationary) ARMA($p_i, p_i - 1$) representation. For instance, a typical trend-cycle decomposition features a random walk (RW) trend ($p_1 = 1$) plus a stationary ARMA(2,1) cycle ($p_2 = 2$). If the parameters are unconstrained, the model is “saturated" and we have to assume a particular value for the correlation of the disturbances driving the components so as to achieve exact identification. If the cycle is specified as a pure AR(2), instead, the reduced form has one more parameter than the UC model and this extra degree of freedom can be used to estimate the correlation between the trend and cycle disturbances.

Morley, Nelson and Zivot (2002, MNZ henceforth) have recently contributed to this issue: they consider a class of UC decompositions of U.S. real gross domestic product (GDP) into a random walk trend and a purely AR(2) cycle, that depends on the identifiable correlation between the trend and cycle disturbances and that produces an ARIMA(2,1,2) reduced form. Within this class, MNZ compare the fit and the components arising from the UC model assuming orthogonal disturbances and the BN decomposition of the unrestricted ARIMA model, which features perfectly and negatively correlated disturbances. The resulting decompositions produce different stylised facts, and in particular the BN cycle is characterised by a much smaller amplitude and a shorter periodicity.

Since a degree of freedom is allowed from the fact that the UC model has one parameter less than the ARIMA reduced form, they estimate the correlation between the trend and cycle disturbances and find out that the estimated value is negative, about -0.92, and significantly different from zero. The resulting *real time*, or concurrent, estimates of the trend and cycle in U.S. GDP closely resemble the BN components, which allows us to reconcile the UC with the unrestricted reduced form.

They interpret this empirical evidence as an expression of the dominant role of real shocks, which shift the long run path of output, whereas short term fluctuations reflect only the adjustment to the new path.

This paper will be concerned with the estimation and the interpretation of decomposition with correlated trend and cycle disturbances. We set up with the specification of the benchmark UC model in Section 2, which nests two leading cases of interest: the UC model with correlated disturbances and the orthogonal decomposition with ARMA(2,1) cyclical component. Section 3 reviews the BN decomposition of the ARIMA(2,1,2) reduced form and its main properties, while Section 4 presents the autocovariance and spectral generating functions of the various models.

In Section 5 we illustrate and compare the fit of orthogonal and correlated UC models to the Euro area and the Italian quarterly real gross domestic product. The evidence for the Italian series mirrors quite closely the findings by MNZ, i.e. in favour of a strong and negative correlation between the trend and cycle disturbances, although we argue that the small sample distribution of the correlation coefficient raises some concern. For the Euro area the findings are not conclusive.

Models with correlated components pose several interpretative issues since, under certain conditions, they are observationally equivalent to models that provide different explanations of the nature of macroeconomic fluctuations (Section 6). Strongly and negatively correlated disturbances imply that the spectral density of the first differences of the series is not a global minimum at the long run frequency. This feature is accommodated also by the cyclical growth model, that can also be parameterised as a model featuring hysteresis effects. The Italian case illustrates that the cyclical growth model and the hysteresis model possess exactly the same explanatory power, yielding the same likelihood.

Inference about unobserved components in models with correlated components is dealt with Section 7, where we consider the state space representation, the treatment of initial conditions, estimation of the components in real time and using the full sample, and some of the ambiguities that arise for single source of error and innovation form representation when they are considered as “models". We also illustrate that a very peculiar trait of models with highly and negatively correlated trend and cycle disturbances is that the future is more informative than the past for signal extraction. As a result the cycle estimates will be subject to large revisions and the final estimates will display greater amplitude than the real time ones.

Sections 9 and 10 investigate respectively whether seasonal adjustment and temporal aggregation can affect the empirical evidence about the sign and the magnitude of the correlation coefficient. Finally, we address the issue as to whether multivariate UC models can cast some light on correlated disturbances (Section 11). Section 12 draws the main conclusions.

2 Trend-Cycle decomposition with Correlated Components

The basic univariate representation for an output series, y_t, deals with the decomposition into a random walk trend component, denoted μ_t, and a stationary ARMA(2,1) stochastic cycle, denoted ψ_t:

$$\begin{array}{rcll} y_t & = & \mu_t + \psi_t & t = 1, 2, \ldots, T, \\ \\ \mu_t & = & \mu_{t-1} + \beta + \eta_t, & \\ \psi_t & = & \phi_1 \psi_{t-1} + \phi_2 \psi_{t-2} + \kappa_t + \theta \kappa_{t-1}, & \\ \left(\begin{array}{c} \eta_t \\ \kappa_t \end{array}\right) & \sim & \text{NID}\left[\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} \sigma_\eta^2 & \sigma_{\eta\kappa} \\ \sigma_{\eta\kappa} & \sigma_\kappa^2 \end{array}\right)\right], & \sigma_{\eta\kappa} = r\sigma_\eta\sigma_\kappa. \end{array} \quad (1)$$

The trend and cycle disturbances are allowed to be contemporaneously correlated, with r being the correlation coefficient; NID denotes normally and independently distributed random variables. Complex stationary autoregressive roots can be imposed expressing $\phi_1 = 2\rho \cos \lambda_c$ and $\phi_2 = -\rho^2$, where ρ and λ_c (representing the modulus and the phase of the roots of the AR characteristic equation), lie respectively in $[0, 1)$ and $[0, \pi]$.

Model (1) will be labelled UC(r, θ) to stress the dependence on the two "conflicting" parameters. Its reduced form is the ARIMA(2,1,2) process:

$$\Delta y_t = \beta + \frac{\theta(L)}{\phi(L)}\xi_t, \quad \xi_t \sim \text{NID}(0, \sigma^2), \qquad t = 2, ..., T, \quad (2)$$

where $\theta(L) = 1 + \theta_1 L + \theta_2 L^2$ and $\phi(L) = 1 - \phi_1 L - \phi_2 L^2$ are respectively the MA and AR polynomials in the lag operator, L, and $\Delta = 1 - L$.

The reduced form has six parameters, whereas UC(r, θ) has seven. Hence, the latter is not identified and one has to restrict either r or θ. The orthogonal trend cycle decomposition considered by Clark (1987) imposes $r = \theta = 0$, and thus will be denoted UC(0,0). MNZ entertain UC$(r, 0)$ and compare it with UC$(0, 0)$. Harvey and Jäger (1994), although they entertain I(2) - local linear - trends, consider UC$(0, \theta)$, with a restricted θ, which functionally depends on ρ and λ_c.

It should be noticed that UC$(r, 0)$ is not identifiable if $\phi_2 = 0$; the parameterisation in terms of the modulus and phase of the AR process avoids this lack of identification if $0 < \rho < 1$.

3 The Beveridge-Nelson Decomposition

The Beveridge and Nelson (1981) decomposition is based upon a definition of the trend in terms of prediction as the value at time t of the eventual forecast function. When y_t is difference stationary, this argument yields a unique decomposition into a random walk trend and a stationary transitory component. The decomposition has been deemed to provide a structural interpretation to any ARIMA$(p, 1, q)$ reduced form model fitted according to the traditional Box-Jenkins methodology, with the characterising property that the components are driven by perfectly correlated disturbances, that are linear in the innovations, ξ_t.

With respect to the ARIMA(2,1,2) model (2), the BN decomposition specialises as follows:

$$y_t = m_t + c_t, \qquad t = 1, ..., T. \quad (3)$$

where the trend, m_t, has the random walk representation:

$$m_t = m_{t-1} + \beta + \frac{\theta(1)}{\phi(1)}\xi_t. \tag{4}$$

and the cycle, c_t, has the ARMA(2,1) representation:

$$\phi(L)c_t = (1 + \vartheta^* L)\left[1 - \frac{\theta(1)}{\phi(1)}\right]\xi_t, \qquad \vartheta^* = -\frac{\phi_2\theta(1) + \theta_2\phi(1)}{\phi(1) - \theta(1)}. \tag{5}$$

These results follow straightforwardly from Proietti (1995) and Proietti and Harvey (2000).

It is apparent from (4) and (5) that the two components are driven by the innovations, ξ_t; the fraction $\theta(1)/\phi(1)$, known as *persistence*, is integrated in the trend, and its complement to 1 drives the cycle. The sign of the correlation between the trend and the cycle disturbances is provided by the sign of $\phi(1) - \theta(1)$; when persistence is less (greater) than one then trend and cycle disturbances are positively (negatively) and perfectly correlated. Furthermore, the BN cycle has always an MA feature, unless $\phi_2\theta(1) + \theta_2\phi(1) = 0$. The MA polynomial can be non invertible, which occurs when $|\vartheta_1^*|$ is greater than 1; this will be the case for the ARIMA(2,1,2) models estimated in Section 5 for the Euro Area and the Italian GDP.

As shown by Watson (1989), the BN components, defined on the reduced form of UC models, are always coincident with the filtered, or real time time, estimates arising from the UC(r, θ) model, whatever restriction we impose to make it identifiable. The filtered components of *identified* UC(r, θ) models are however estimated with non zero mean square error even in the case $r = -1, \theta = 0$. Hence, it would not be correct to regard the BN trend and cycle as the estimates of the components arising from UC(-1,0), as future observations reduce the estimation error. This point has often been overlooked in the literature and we return to it in Section 8, where we show that the only case in which ψ_t is actually an observed component in real time arises for $r = 1$.

When the BN decomposition is interpreted *as a model*, the components are estimated in real time with zero mean square error, after processing a suitable number of observations so that the effect of initial conditions is marginalised, this being the only source of uncertainty (assuming known parameters). We discuss this further in Section 7.

4 Autocovariance Generating Functions

The properties of any linear time series model of economic fluctuations are uniquely characterised by its autocovariance generating function (ACGF). The ACGF also provides a valuable tool to address the equivalence issues that arise in the interpretation of models with correlated disturbances (see Section 6). Moreover, its frequency domain counterpart, the spectral generating function (SGF), will be used for estimating the parameters of the model by maximum likelihood.

The ACGF of the reduced form model for Δy_t, denoted $g(L)$, is $g(L) = \sigma^2|\theta(L)|^2/|\phi(L)|^2$, where $|\theta(L)|^2 = \theta(L)\theta(L^{-1})$ and $|\phi(L)|^2 = \phi(L)\phi(L^{-1})$. The various UC model that result by constraining (1) are restricted versions of $g(L)$. For the UC(r,0) model considered by MNZ the ACGF, denoted $g_r(L)$, can be written as follows (Proietti, 2002):

$$|\phi(L)|^2 g_r(L) = |\phi(L)|^2\sigma_\eta^2 + |1-L|^2[\sigma_\kappa^2 + r\sigma_\eta\sigma_\kappa(1+\phi_1+\phi_2+\phi_2(L+L^{-1}))]. \qquad (6)$$

Equating $g(L)$ to $g_r(L)$ provides the way of deriving the reduced form parameters $(\theta_1, \theta_2, \sigma^2)$ from $(\sigma_\eta, \sigma_\kappa, r)$ and of assessing the restrictions imposed by the UC model on the reduced form. For instance, $g(1) = g_r(1)$ implies $\sigma_\eta^2 = \sigma^2[\theta(1)/\phi(1)]^2$.

For the UC$(0,\theta)$ model we have

$$|\phi(L)|^2 g_\theta(L) = |\phi(L)|^2\sigma_{\eta^*}^2 + |1-L|^2|1+\theta L|^2\sigma_{\kappa^*}^2 \qquad (7)$$

where, with a change of notation that will be useful in the sequel, $\sigma_{\eta^*}^2$ and $\sigma_{\kappa^*}^2$ denote the variance of the trend and cycle disturbances when we assume in (1) that they are mutually uncorrelated at all leads and lags.

The ACGF of the Clark model, UC(0,0), is obtained by setting $r = 0$ in (6) or $\theta = 0$ in (7). Replacing L with the complex exponential $e^{-\imath\lambda} = \cos\lambda - \imath\sin\lambda$, where $\imath$ is the imaginary unit, gives the spectral generating function, that provides a decomposition of the variance of Δy_t into the contribution of changes in the trend, in the cycle and, in the case of UC(r,0), the covariation (cross spectral density) between the two.

5 Two Illustrative Examples

This section illustrates the fit of the unrestricted ARIMA(2,1,2) and three different trend-cycle decompositions that result from (1), UC(0,0), UC$(r,0)$ and UC$(0,\theta)$, with respect to Euro Area (EA) and the Italian Gross Domestic Product (GDP) at constant prices. Both series are quarterly and are available for the sample period 1970:1-2002.2. The EA series is an update of the one constructed for the Area Wide Model by Fagan, Henry and Mestre (2001), and the Italian series is made available electronically at www.istat.it.

Model estimation has been carried out in the frequency domain. The likelihood is defined in terms of the stationary representation of the various models, that is in terms of $\Delta y_t, t = 1, \dots, T^* = (T-1)$; see Nerlove, Grether and Carvalho (1995) and Harvey (1989, sec. 4.3).

While the time domain likelihood of UC models is based on a recursive orthogonalisation, known as the prediction error decomposition, performed by the Kalman filter (Section 7), the frequency domain one is based on an alternative orthogonalisation, achieved through a Fourier transform.

Denoting the Fourier frequencies by $\lambda_j = \frac{2\pi j}{T^*}$, $j = 0, 1, \dots, (T^*-1)$, the likelihood function is defined as follows:

$$\text{loglik} = -\frac{1}{2}\left\{T^*\log 2\pi + \sum_{j=0}^{T^*-1}\left[\log g_m(\lambda_j) + 2\pi\frac{I(\lambda_j)}{g_m(\lambda_j)}\right]\right\}$$

where $g_m(\lambda_j) = g_m(e^{-\imath\lambda_j})$ denote the spectral generating function of the m-th model evaluated at frequency λ_j, and $I(\lambda_j)$ is the periodogram:

$$I(\lambda_j) = \frac{1}{2\pi}\left[c_0 + 2\sum_{\tau=1}^{T^*-1} c_\tau \cos(\lambda_j \tau)\right]$$

where c_τ denotes the sample autocovariance at lag τ,

$$c_\tau = \frac{1}{T^*}\sum_{t=1}^{T-\tau}(\Delta y_t - \bar{\Delta y})(\Delta y_{t-\tau} - \bar{\Delta y}), \quad \bar{\Delta y} = \frac{1}{T^*}\sum_{t=1}^{T^*}\Delta y_t.$$

The index m refers alternatively to the ARIMA model, UC(0,0), UC$(r,0)$, and UC$(0,\theta)$. The corresponding spectral generating functions are straightforwardly derived from the ACGFs presented in Section 4[2].

Table 1 presents the main estimation results along with some diagnostics: $Q(12)$ denotes the Ljung-Box portmanteau test statistic for residual autocorrelation based on the first 12 autocorrelations, and we also present the Doornik and Hansen (1994) test of normality. Both are computed on the standardised Kalman filter innovations (see appendix 7).

We observe that for both series the ARIMA model and UC$(r,0)$ provide exactly the same likelihood inferences; hence the reduced form of the latter coincides with the unrestricted ARIMA(2,1,2) model fitted to the series. The persistence parameter is respectively 1.38 (EA) and 1.13 (Italy). The estimated correlation parameter is high and negative (-0.95 for EA and -0.82 for Italy), and the ratio $\sigma_\eta/\sigma_\kappa$ is always greater that 1. The likelihood ratio (LR) test of the restriction $r = 0$ is not significant for EA but it is highly so for Italy, with a p-value 0.02, whereas the LR test of $\theta = 0$ is never significant. It is noticeable that for Italy the estimated cycle MA parameter lies on the boundary of the parameter space; in general models with orthogonal disturbances yield worse Ljung-Box statistics.

Table 1: Parameter estimates and diagnostics for models of quarterly Euro Area and Italian GDP, 1970.1-2002.2; (r) denotes a restricted parameter

Estimates	Euro Area				Italy			
	ARIMA	UC$(0,0)$	UC$(r,0)$	UC$(0,\theta)$	ARIMA	UC$(0,0)$	UC$(r,0)$	UC$(0,\theta)$
ϕ_1	1.40	1.65	1.40	1.65	1.47	1.54	1.47	1.56
ϕ_2	-0.69	-0.68	-0.69	-0.68	-0.77	-0.59	-0.77	-0.84
θ_1	-1.17				-1.14			
θ_2	0.57				0.48			
σ^2	0.3443				0.5225			
r		0(r)	-0.95	0(r)		0(r)	-0.82	0(r)
σ_η^2 $(\sigma_{\eta^*}^2)$		0.2090	0.6473	0.2089		0.1475	0.6672	0.3957
σ_κ^2 $(\sigma_{\kappa^*}^2)$		0.0975	0.2226	0.0975		0.3560	0.2539	0.0216
θ		0(r)	0(r)	-0.02		0(r)	0(r)	1.00
loglik	-114.28	-114.52	-114.28	-114.51	-141.17	-144.04	-141.17	-143.03
$Q(12)$	6.83	7.70	6.83	7.76	6.29	13.53	6.29	14.07
Normality	11.80	11.63	11.80	11.58	2.08	1.53	2.08	2.46

[2] All the computations were performed in Ox 3.2 (Doornik, 2001). Signal extraction was performed by the Kalman filter and smoother using the library of state space function SsfPack 3.0 (beta) by Koopman et al. (1999), linked to Ox 3.2.

The fit provided by the models to the periodogram (raw sample spectrum) emerges from Figure 1, which presents $I(\lambda_j)$ along with the estimated spectral density functions $g_r(\lambda_j)/(2\pi)$ and $g_\theta(\lambda_j)/(2\pi)$. Obviously, $g_r(\lambda_j) = g(\lambda_j)$, that is the ARIMA spectrum is identical to that implied by UC(r,0).

For EA the spectral density fitted by UC(0,θ) is characterised by a spectral peak taking place at a lower frequency, and therefore the resulting cycle estimates are characterised by a larger period (the estimated AR polynomial can actually be written as $\phi(L) = (1 - 0.82L)^2$ and thus features a stationary root with multiplicity 2 at the zero frequency - this is a special case of a second order cycle, see Harvey and Trimbur, 2002); as expected, UC(r, 0), yields a much higher estimate at the zero frequency and it is not a minimum at that frequency. We also notice a periodogram ordinate close to the Nyquist frequency that is not fitted by any of the models: this is the likely effect of an additive outlier occurring in 1974.4, which also inflates the normality test statistic.

For Italy, the cycle periods are not different (about 3 years), although $g_r(\lambda_j)$ and $g_\theta(\lambda_j)$ differ around the zero frequency and the spectral peak. For the latter we observe that $\theta = 1$ implies $g_\theta(0) = g_\theta(\pi)$ as the cycle is strictly non invertible at the π frequency. The richer residual autocorrelation pattern characterising UC(0,θ) are likely to be a consequence of underestimation of the zero frequency variance component.

For estimation purposes, we adopted the transformation $r = \bar{r}/\sqrt{1+\bar{r}^2}$, where $\bar{r}$ is estimated unrestrictedly and the transformation ensures that the correlation parameter is constrained in the admissible range [-1,1]. The asymptotic standard error of r, estimated by the Delta method, are 0.17 and 0.22, respectively for the EA and Italy. These, however, provide only a very bad guidance over the sampling distribution of r. To illustrate this point we generated 1000 bootstrap estimates of r; the distribution is plotted in Figure 2. For the implementation of the bootstrap we followed Stoffer and Wall (1991), generating 1000 series with the same sample size of the original ones by resampling without replacement the standardised innovations arising from the fitted UC(r,0) model. The sampling distribution is highly nonstandard as it suffers from a "pile-up" phenomenon at the extremes of the sample range; if we consider that in the Italian case 27% and 6% are equal respectively to -1 and +1, a bootstrap confidence interval covers all the parameter range. The same considerations apply to EA.

In conclusion, the results presented in this section confirm the MNZ findings, pointing out that among the unobserved components models considered, the UC(r, 0) model is the only one that can be reconciled with the unrestricted ARIMA(2,1,2) model of GDP. However, only for Italy the correlation between trend and cycle disturbances resulted significant using standard asymptotic inferences. The bootstrap characterisation of the sampling distribution of the correlation parameter suggests that those inferences need to be handled with great care.

6 Interpretative Issues

The evidence emerging from our empirical illustrations, would, with the caveats made above, point out in the direction of selecting the UC(r, 0) specification, with a high and negative correlation between the trend and cycle disturbances. This, along with the signal ratio $\sigma^2_\eta/\sigma^2_\kappa$

being relatively high, has been taken to support the notion of the prominence of real shocks, η_t, as opposed to nominal ones, κ_t.

In this section we review some alternative ways of interpreting the correlation between the disturbances, by establishing the conditions under which UC(r, 0) can be viewed as a reparameterisation of an alternative decomposition with a very different meaning.

6.1 The Equivalence of UC(r,0) and UC($0, \theta$)

The UC(r,0) model can be rewritten as an UC(0,θ) model if the quadratic equation,

$$\frac{\theta}{1+\theta^2} = \phi_2 r \frac{\sigma_\eta}{\sigma_\kappa}\left[1 + r\frac{\sigma_\eta}{\sigma_\kappa}(1+\phi_1+\phi_2)\right]^{-1}, \tag{8}$$

admits a real and invertible solution (Proietti, 2002). The remaining parameters are then obtained as follows: $\sigma^2_{\eta^*} = \sigma^2_\eta$ and $\sigma^2_{\kappa^*} = \phi_2 r \sigma_\eta \sigma_\kappa/\theta$. These results are derived form the ACGF identity, $g_\theta(L) = g_r(L)$, which amounts to equating the right hand sides in (6) and (7).

The admissibility conditions can be shown to be exactly the same under which $g_r(e^{-\imath\lambda})$ is a global minimum at the zero frequency. This reflects the fundamental fact that the orthogonal decomposition UC($0, \theta$) imposes that the spectral density of Δy_t is a minimum at zero, a result already established in Lippi and Reichlin (1992).

The equivalence is always feasible if r is positive, but, we can allow for negative correlation provided that the ratio $\sigma_\eta/\sigma_\kappa$ is small, i.e. the trend disturbance is a minor source of variation. In conclusion, when the spectral density of Δy_t is a minimum at zero, the cross spectrum between the components absorbs part of the cyclical variability; this can be reallocated to the cyclical component, which is underestimated by the UC(r,0) model, by allowing it to display a moving average feature.

6.2 Cyclical Growth and Hysteresis

Consider now the following UC model that postulates that Δy_t can be additively decomposed into a cyclical component and orthogonal noise:

$$\begin{array}{rcll} \Delta y_t & = & \beta + \psi_t + \eta^*_t, & \eta^*_t \sim \mathrm{WN}(0, \sigma^2_{\eta^*}), \\ \psi_t & = & \phi_1\psi_{t-1} + \phi_2\psi_{t-2} + \kappa^*_t + \theta\kappa^*_{t-1}, & \kappa^*_t \sim \mathrm{WN}(0, \sigma^2_{\kappa^*}), \\ \mathrm{E}(\eta^*_t\kappa^*_t) & = & 0. & \end{array} \tag{9}$$

The idea is that of representing underlying growth as a smooth cyclical process.

Model (9) has again an ARIMA(2,1,2) reduced form, and six parameters, but different implications. In its original specification, it simply produces estimates of underlying growth that are smoother than the original observations; it can also be interpreted as a *cyclical trend model*, as in Harvey (1989, p. 46), such that the trend is coincident with the observations, i.e. $y_t = \mu_t$ and $\mu_t = \mu_{t-1} + \beta + \psi_t + \eta^*_t$.

It is also observationally equivalent to the Jäger and Parkinson (1994) *hysteresis* model, which is such that a deviation cycle can still be defined, but the cycle modifies also permanently the trend. The hysteresis model is specified as follows:

$$\begin{aligned}
y_t &= \mu_t + \psi_t^*, & t = 1, 2, \ldots, T, \\
\mu_t &= \mu_{t-1} + (1+\theta)\psi_{t-1}^* + \eta_t^*, & \eta_t^* \sim \mathrm{WN}(0, \sigma_{\eta^*}^2), \\
\psi_t^* &= \phi_1 \psi_{t-1}^* + \phi_2 \psi_{t-2}^* + \kappa_t^*, & \kappa_t^* \sim \mathrm{WN}(0, \sigma_{\kappa^*}^2)
\end{aligned} \tag{10}$$

and $\mathrm{E}(\eta_t^* \kappa_t^*) = 0$. Notice that the cycle, ψ_t^*, is redefined as a pure second order AR process; $(1+\theta)$ represents the hysteresis parameter, i.e. the fraction of the cycle that is integrated in the trend. Obviously, $\theta = -1$ yields again the additive decomposition into orthogonal trend and cycle that corresponds to the Clark model UC(0, 0).

Using the same expedient of equating the ACGFs, we establish a set of conditions under which (9) can provide a trend - cycle decomposition with correlated disturbances, i.e. can be written as an UC(r, 0) process. These are met if we can uniquely determine the cycle MA parameter θ in (9) for given values of the correlation parameter r and the ratio $\sigma_\eta/\sigma_\kappa$ in UC(r, 0), as the admissible invertible solution of the quadratic equation:

$$\frac{(1+\theta)^2}{(1+\theta)^2\left[\phi_1(1-\phi_2) + 2\phi_2\right] + \theta\phi(1)^2} = \frac{r(\sigma_\eta/\sigma_\kappa)}{1 + r(\sigma_\eta/\sigma_\kappa)(1+\phi_1+\phi_2)}. \tag{11}$$

If this is possible, then, the remaining parameters are obtained from:

$$\sigma_{\kappa^*}^2 = -r\sigma_\eta\sigma_\kappa \frac{\phi(1)^2}{(1+\theta)^2}; \qquad \sigma_{\eta^*}^2 = \sigma_\eta^2 - \frac{(1+\theta)^2}{\phi(1)^2}\sigma_{\kappa^*}^2.$$

These results make clear that the equivalence is admissible only for negative values of r. When $r = 0$ the solution $\theta = -1$ arises for any value of the ratio $\sigma_\eta/\sigma_\kappa$, in which case the hysteresis parameter is zero and the model can be orthogonally decomposed into a RW trend and a purely AR(2) cycle. No admissible solutions exists for a positive r and in general an UC trend-cycle decomposition with positively correlated disturbances cannot be isomorphic to a cyclical growth model or a model with hysteresis effects. This is so since model (9) implies a spectral density for Δy_t that has a local, but not a global, minimum at the zero frequency.

6.3 Illustrative Examples (cont.)

The parameter estimates reported in Table 1 rule out the equivalence of the estimated UC(r, 0) model with UC(0,θ): for both EA and Italy 8 has no admissible solution. As a matter of fact, the spectral density estimated by the former is not a global minimum at the zero frequency.

As far as the equivalence with the cyclical growth - hysteresis model is concerned, for the Euro Area GDP equation (11) has complex roots. However, for the Italian GDP case a real invertible solution is admissible as $\theta = -0.41$; the remaining implied parameter values are $\sigma_{\eta*}^2 = 0.3260$ and $\sigma_{\kappa*}^2 = 0.0869$; These values are fully coincident with those estimated by maximum likelihood. Hence the Italian GDP provides a case in which the cyclical growth model and the trend cycle decomposition with correlated disturbances provide exactly the same inferences, that are in turn coincident with those arising for the unrestricted ARIMA(2,1,2) model. As a result, alternative explanations of the nature of macroeconomic fluctuations arise with exactly the same likelihood.

7 State Space Representation and the Estimation of Unobserved Components

In this section we discuss several facts concerning the real time and smoothed estimates of trends and cycles arising from the various models considered in the previous sections. To accomplish this, we need first to review the state space representation, and the associated algorithms for filtering and smoothing.

7.1 State Space representation

The UC models considered so far admit the time-invariant state space representation:

$$\begin{array}{rcll} y_t &=& \boldsymbol{z}'\boldsymbol{\alpha}_t, & t = 1, 2, \dots, T, \\ \boldsymbol{\alpha}_t &=& \boldsymbol{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{c} + \boldsymbol{R}\boldsymbol{\epsilon}_t, & \end{array} \tag{12}$$

with $\boldsymbol{\epsilon}_t \sim \text{NID}(\mathbf{0}, \boldsymbol{Q})$ and $\boldsymbol{\alpha}_0 \sim \text{NID}(\tilde{\boldsymbol{\alpha}}_0, \boldsymbol{P}_0)$, independently of $\boldsymbol{\epsilon}_t, \forall t$. The treatment of initial conditions is discussed in Section 7.2. The state vector has three elements, $\boldsymbol{\alpha}_t = [\mu_t, \psi_t, \psi_t^*]$, and the drift β is considered as a constant effect. Alternatively, we may include β in the state vector using the transition equations for the trend $\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$, $\beta_t = \beta_{t-1}$.

For instance, the system matrices for the UC$(r, 0)$ model are:

$$\boldsymbol{z} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & 1 \\ 0 & \phi_2 & 0 \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} \beta \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{Q} = \begin{bmatrix} \sigma_\eta^2 & \sigma_{\eta\kappa} \\ \sigma_{\eta\kappa} & \sigma_\kappa^2 \end{bmatrix}$$

For the Beveridge-Nelson decomposition, considered *as a model*, the system matrices are the same except for $\boldsymbol{R}$ and $\boldsymbol{Q}$, which are 3×1 and scalar, respectively:

$$\boldsymbol{R} = \begin{bmatrix} \varrho \\ 1 - \varrho \\ -(\theta_2 + \phi_2\varrho) \end{bmatrix}, \quad \boldsymbol{Q} = \sigma^2,$$

where $\varrho = \theta(1)/\phi(1)$ is the persistence parameter. On the other hand, for UC(0,θ) we need to replace $\boldsymbol{R}$ and $\boldsymbol{Q}$ by:

$$\boldsymbol{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & \theta \end{bmatrix}, \quad \boldsymbol{Q} = \begin{bmatrix} \sigma_{\eta^*}^2 & 0 \\ 0 & \sigma_{\kappa^*}^2 \end{bmatrix};$$

finally, the state space representation for the cyclical growth model is obtained also replacing $\boldsymbol{z}$ and $\boldsymbol{T}$ by:

$$\boldsymbol{z} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{T} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & \phi_1 & 1 \\ 0 & \phi_2 & 0 \end{bmatrix}.$$

The state space model (12) is in *contemporaneous form*. The *future form*, that is sometimes used to specify the model, differs for the timing of the transition equation, which is written $\boldsymbol{\alpha}_{t+1} = \boldsymbol{T\alpha}_t + \boldsymbol{c} + \boldsymbol{R\epsilon}_t$. Due to the absence of measurement noise and time invariance of the system matrices, if the measurement equation is unaltered, that is $y_t = \boldsymbol{z}'\boldsymbol{\alpha}_t$, the two representations differ only for the (arbitrary) timing of the disturbances. A different representation is obtained if the maintained model is in contemporaneous form and we express it to the future form: this can be done by replacing $\boldsymbol{\alpha}_t$ in the measurement equation with the right hand side of the transition equation and redefining $\boldsymbol{\alpha}_t^* = \boldsymbol{\alpha}_{t-1}$, so as to write:

$$\begin{aligned} y_t &= \boldsymbol{z}'\boldsymbol{T\alpha}_t^* + \boldsymbol{z}'\boldsymbol{c} + \boldsymbol{z}'\boldsymbol{R\epsilon}_t, \qquad t = 1, 2, \dots, T, \\ \boldsymbol{\alpha}_{t+1}^* &= \boldsymbol{T\alpha}_t^* + \boldsymbol{c} + \boldsymbol{R\epsilon}_t. \end{aligned} \tag{13}$$

It should be noticed the appearance of measurement noise in the first equation that is correlated with the transition noise. There is nothing “structural" about this component, which appears as a consequence of the operation of forcing the contemporaneous representation into the future form, according to which a “shock" at time t affects the components at time $t+1$.

7.2 Initialisation

Initialisation deals with the specification of the mean and the covariance matrix of the initial state vector, $\boldsymbol{\alpha}_0$. If we assume that the process $\boldsymbol{\alpha}_t$ has applied since time immemorial ($t \to -\infty$), and if we partition $\boldsymbol{T} = \text{diag}(1, \boldsymbol{T}_\psi)$ and $\boldsymbol{R} = [\boldsymbol{R}_\mu', \boldsymbol{R}_\psi']'$, then $\tilde{\boldsymbol{a}}_0 = \boldsymbol{0}$ and

$$\boldsymbol{P}_0 = \boldsymbol{e}_1\boldsymbol{e}_1'\delta + \begin{bmatrix} 0 & \boldsymbol{d} \\ \boldsymbol{d}' & \boldsymbol{M} \end{bmatrix}$$

where $\delta \to \infty$, $\boldsymbol{e}_1' = [1, 0, 0]$, $\boldsymbol{d} = \boldsymbol{R}_\mu\boldsymbol{Q}\boldsymbol{R}_\psi'(\boldsymbol{I} - \boldsymbol{T}_\psi')^{-1}$, and $\boldsymbol{M}$ solves the matrix equation $\boldsymbol{M} = \boldsymbol{T}_\psi\boldsymbol{M}\boldsymbol{T}_\psi' + \boldsymbol{R}_\psi\boldsymbol{Q}\boldsymbol{R}_\psi'$. UC models with uncorrelated components have $\boldsymbol{d} = \boldsymbol{0}'$, whereas, for UC($r$,0), $\boldsymbol{d} = r\sigma_\eta\sigma_\kappa\phi(1)^{-1} \cdot [1, \ \phi_2]$. However, working out the exact initial KF by letting $\delta \to \infty$, as in Koopman (1997) and Durbin and Koopman (2001), it can be checked that the elements of $\boldsymbol{d}$ are wiped away by the limiting operations and thus play no role for inferences. The same result is obtained if one uses the augmentation approach, and in particular the theory in De Jong and Chu-Chun-Lin (1994).

7.3 Kalman Filter

The Kalman filter (Anderson and Moore, 1979), is the well-known recursive algorithm for computing the minimum mean square estimator of $\boldsymbol{\alpha}_t$ and its mean square error (MSE) matrix conditional on $Y_{t-1} = \{y_1, y_2, \dots, y_{t-1}\}$. Defining $\tilde{\boldsymbol{\alpha}}_{t|t-1} = \text{E}(\boldsymbol{\alpha}_t|Y_{t-1})$, $\boldsymbol{P}_{t|t-1} = \text{E}[(\boldsymbol{\alpha}_t -$

$\tilde{\boldsymbol{\alpha}}_{t|t-1})(\boldsymbol{\alpha}_t - \tilde{\boldsymbol{\alpha}}_{t|t-1})'|Y_{t-1}]$, it is given by the set of recursions:

$$\begin{array}{rclrcl} \xi_t & = & y_t - \boldsymbol{z}'\tilde{\boldsymbol{\alpha}}_{t|t-1}, & f_t & = & \boldsymbol{z}'\boldsymbol{P}_{t|t-1}\boldsymbol{z} \\ & & & \boldsymbol{k}_t & = & \boldsymbol{T}\boldsymbol{P}_{t|t-1}\boldsymbol{z}f_t^{-1} \\ \tilde{\boldsymbol{\alpha}}_{t+1|t} & = & \boldsymbol{T}\tilde{\boldsymbol{\alpha}}_{t|t-1} + \boldsymbol{c} + \boldsymbol{k}_t\xi_t, & \boldsymbol{P}_{t+1|t} & = & \boldsymbol{T}\boldsymbol{P}_{t|t-1}\boldsymbol{T}' + \boldsymbol{R}\boldsymbol{Q}\boldsymbol{R}' - \boldsymbol{k}_t\boldsymbol{k}_t' f_t \end{array} \qquad (14)$$

$\xi_t = y_t - \mathrm{E}(y_t|Y_{t-1})$ are the filter innovations or one-step-ahead prediction errors, with variance f_t.

Steady State The innovations and the state one-step-ahead prediction error, $\boldsymbol{x}_t = \boldsymbol{\alpha}_t - \tilde{\boldsymbol{\alpha}}_{t|t-1}$, can be written as

$$\xi_t = \boldsymbol{z}'\boldsymbol{x}_t, \quad \boldsymbol{x}_{t+1} = \boldsymbol{L}_t\boldsymbol{x}_t + \boldsymbol{R}\boldsymbol{\epsilon}_{t+1}, \qquad (15)$$

where $\boldsymbol{L}_t = \boldsymbol{T} - \boldsymbol{k}_t\boldsymbol{z}'$. Thus, $\boldsymbol{x}_t$ follows a VAR(1) process that is (asymptotically) stationary if the autoregressive matrix $\boldsymbol{L}_t$, known as the *closed loop matrix* in system theory, converges to a matrix $\boldsymbol{L} = \boldsymbol{T} - \boldsymbol{k}\boldsymbol{z}'$, whose eigenvalues lie all inside the unit circle.

The basic properties that ensure convergence to such stabilising solution are *detectability* and *stabilisability* (see Burridge and Wallis, 1988). For the trend-cycle decompositions considered in this paper they are met if $\phi(L)$ does not display explosive roots or unit roots at the zero frequency. The two conditions imply that, independently of initial conditions, $\boldsymbol{P}_{t+1|t}$ converges at an exponential rate to a steady state solution $\boldsymbol{P}$, satisfying the Riccati equation $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{P}\boldsymbol{T}' + \boldsymbol{R}\boldsymbol{Q}\boldsymbol{R}' - \boldsymbol{k}\boldsymbol{k}'f$, with $\boldsymbol{k} = \boldsymbol{T}\boldsymbol{P}\boldsymbol{z}f^{-1}$ and $f = \boldsymbol{z}'\boldsymbol{P}\boldsymbol{z}$, and the Kalman gain vector $\boldsymbol{k}$ is such that $\boldsymbol{L}$ has all its eigenvalues inside the unit circle.

7.4 Real time estimates

The real time or concurrent estimates of the states and the estimation error covariance matrix are given respectively by:

$$\tilde{\boldsymbol{\alpha}}_{t|t} = \tilde{\boldsymbol{\alpha}}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{z}f_t^{-1}\xi_t, \quad \boldsymbol{P}_{t|t} = \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{z}\boldsymbol{z}'\boldsymbol{P}_{t|t-1}f_t^{-1}. \qquad (16)$$

The estimated unobserved components in $\tilde{\boldsymbol{\alpha}}_{t|t}$ are the same as those arising from the BN decomposition of the implied ARIMA reduced form representation. The MA parameters of the reduced form representation can be uniquely derived from the steady state using $\boldsymbol{P}\boldsymbol{z}f^{-1}$, whose first element is the persistence parameter. Notice, however, that in the steady state we need $\boldsymbol{z}\boldsymbol{z}'f^{-1}$ to be equal to the pseudo-inverse of $\boldsymbol{P}$ for the components to be estimated with zero error, i.e. observable with respect to current and past information. For the BN model $f = \sigma^2 = \boldsymbol{Q}$, $\boldsymbol{P}\boldsymbol{z}f^{-1} = \boldsymbol{R}$ and $\boldsymbol{P} = \boldsymbol{R}\boldsymbol{Q}\boldsymbol{R}'$, $\boldsymbol{k} = \boldsymbol{T}\boldsymbol{R}$, which ensures that when the system has reached a steady state, the components are estimated in real time with zero mean square error.

7.5 Single Source of Error and Innovation State Space Models

The BN decomposition can be viewed as the "structural" representation of *Single Source of Error* (SSE, Snyder, 1985, Hyndman *et al.*, 2002) and *Steady State Innovation Models* (SSIM, Casals, Jerez and Sotoca, 2002). In this section we establish the connection among these alternative representations of the same underlying model.

Recalling the state space representation of the BN decomposition, which is (12) with scalar $\boldsymbol{\epsilon}_t = \xi_t$ and $\boldsymbol{R} = [\varrho, (1-\varrho), -(\theta_2 + \phi_2\varrho)]$, we use manipulations similar to those which led to (13) in two steps: we first replace $\boldsymbol{\alpha}_t$ in the measurement equation by the right hand side of the transition equation to obtain the single source of error representation:

$$y_t = \boldsymbol{z}'\boldsymbol{T}\boldsymbol{\alpha}_{t-1} + \xi_t, \quad \boldsymbol{\alpha}_t = \boldsymbol{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{c} + \boldsymbol{R}\xi_t,$$

the new measurement equation features ξ_t since $\boldsymbol{z}'\boldsymbol{R} = \boldsymbol{z}'\boldsymbol{r} = 1$.

Next, we posit $\boldsymbol{\alpha}_t^* = \boldsymbol{T}\boldsymbol{\alpha}_{t-1}$ and on premultiplying both sides of the transition equation by $\boldsymbol{T}$, we write:

$$\begin{array}{rcll} y_t & = & \boldsymbol{z}'\boldsymbol{\alpha}_t^* + \xi_t, & t = 1, 2, \ldots, T, \\ \boldsymbol{\alpha}_{t+1}^* & = & \boldsymbol{T}\boldsymbol{\alpha}_t^* + \boldsymbol{c} + \boldsymbol{k}\xi_t, & \end{array} \qquad (17)$$

with $\boldsymbol{k} = \boldsymbol{T}\boldsymbol{R}$. This is the innovation form of the model in the steady state, as can be seen by comparing (17) with the first and the last row of the KF equations in (14). The Kalman gain identity is $\boldsymbol{k} = \boldsymbol{T}\boldsymbol{R}$.

We stress that both forms are available for any UC model, and in fact the SSIM arises from the filtering operation, and the SSE from the updating equations, but only for the BN model a one to one correspondence holds. If SSIM or SSE are estimated as a model there is no way of recovering information on multiple source of errors.

The emphasis on the “exact" nature of inference for SSE and SSIM models is misplaced: the property that components are estimated in real time, state Casals, Jerez and Sotoca (2002) in their concluding remarks (p. 563),

> “ensures coherence between the properties of the theoretical and empirical components, provides a rigorous statistical foundation for using the empirical components as observable and mutually independent time series, and guarantees that these components will not change as the sample increases".

First and foremost, the BN illustration shows that both SSIM and SSE representations feature (perfectly correlated) measurement noise that is not present in the BN model, and thus are not *coherent* in this respect with the maintained model. As a matter of fact, Casals, Jerez and Sotoca interpret the one-step-ahead forecast error as an estimate of the irregular component. Moreover, the timing of the disturbances is not the same as the original model, when the latter is expressed in contemporaneous form. Secondly, the empirical components cannot be mutually independent, being driven by the same disturbance. Finally, their observability in real time, and thus the absence of revision is simply a consequence of the model formulation, i.e. a property and not necessarily an advantage.

We do not need much theory to show that when (17) is interpreted as a model, and the system is stabilisable and detectable, the states are observed: substituting the expression for ξ_t in the measurement equation into the transition equation, yields $\boldsymbol{\alpha}_{t+1}^* = \boldsymbol{L}\boldsymbol{\alpha}_t^* + \boldsymbol{c} + \boldsymbol{k}y_t$, that is

$$\boldsymbol{\alpha}_{t+1}^* = \boldsymbol{k}y_t + \boldsymbol{L}\boldsymbol{k}y_{t-1} + \boldsymbol{L}^2\boldsymbol{k}y_{t-2} + \cdots .$$

This expression makes it apparent that the states are a linear combination of *past* observations. This is so since the “states" are in fact one-step-ahead predictions.

In conclusions, SSIM and SSE are useful for prediction, but when they are used for estimation of unobserved components they are prone to a number of inconsistencies and a variety of interpretative issues.

7.6 Smoothing and Final Estimates

We can keep track of revisions, due to the accrual of further observations, by using a *fixed-point smoothing* algorithm. Elaborating results in de Jong (1989), and assuming that the system has reached a steady state, we have, for a fixed t and for $l \geq 0$, the following smoothing recursions:

$$\begin{array}{rclrcl} \tilde{\boldsymbol{\alpha}}_{t|t+l} & = & \tilde{\boldsymbol{\alpha}}_{t|t} + \boldsymbol{P}\boldsymbol{L}'\boldsymbol{r}_{t|t+l}, & \boldsymbol{P}_{t|t+l} & = & \bar{\boldsymbol{P}}_{t|t} - \boldsymbol{P}\boldsymbol{L}'\boldsymbol{N}_{t|t+l}\boldsymbol{L}\boldsymbol{P}, \\ \boldsymbol{r}_{j|t+l} & = & \boldsymbol{L}'\boldsymbol{r}_{j+1|t+l} + \boldsymbol{z}f^{-1}\xi_{j+1}, & \boldsymbol{N}_{j|t+l} & = & \boldsymbol{L}'\boldsymbol{N}_{j+1|t+l}\boldsymbol{L} + \boldsymbol{z}\boldsymbol{z}'f^{-1}, \end{array} \tag{18}$$

$j = t+l, t+l-1, \ldots, t$, where $\bar{\boldsymbol{P}}_{t|t} = \boldsymbol{P} - \boldsymbol{P}\boldsymbol{z}\boldsymbol{z}'\boldsymbol{P}f^{-1}$ and the backwards recursions are initialised $\boldsymbol{r}_{t+l|t+l} = \boldsymbol{0}$, $\boldsymbol{N}_{t+l|t+l} = \boldsymbol{0}$.

Now, as $l \to \infty$ (i.e. assuming a doubly infinite sample), $\boldsymbol{r}_{j|t+l}$ is a backward first order stationary vector autoregression, and $\boldsymbol{N}_{j|t+l}$ is its covariance matrix. The final state estimation error covariance matrix, denoted $\boldsymbol{P}_{t|\infty}$, solves $\boldsymbol{P}_{t|\infty} = \boldsymbol{P} - \boldsymbol{P}\boldsymbol{N}\boldsymbol{P}$, where $\boldsymbol{N}$ is the steady state solution of the backward smoothing equation, $\boldsymbol{N}_{j|t+l} = \boldsymbol{L}'\boldsymbol{N}_{j+1|t+l}\boldsymbol{L} + \boldsymbol{z}\boldsymbol{z}'f^{-1}, j = t+l, \ldots, t$, as $l \to \infty$; a unique stable solution for $\boldsymbol{N}$ exists provided the characteristic roots of $\boldsymbol{L}$ are less than unity in modulus, which is already the condition for a steady state solution. The elements of the solution are obtained from

$$\text{vec}(\boldsymbol{N}) = (\boldsymbol{I} - \boldsymbol{L}' \otimes \boldsymbol{L}')^{-1}\text{vec}(\boldsymbol{z}\boldsymbol{z}'f^{-1}).$$

Hence, $\boldsymbol{P}_{t|\infty}$ contains the final estimation error covariance matrix, and can be written:

$$\boldsymbol{P}_{t|\infty} = \bar{\boldsymbol{P}}_{t|t} - \boldsymbol{P}(\boldsymbol{z}\boldsymbol{z}'f^{-1} - \boldsymbol{N})\boldsymbol{P}.$$

The second term on the right hand side, which is obviously positive semi-definite, measures the total reduction in the estimation uncertainty as we go from the real time to the final estimates.

8 Illustrative Examples (cont.)

Figure 3 displays the smoothed estimates of the components of the EA GDP arising from the UC(r, 0) and UC(0,θ) models estimated in Section 5. We also present the real time estimates, $\tilde{\psi}_{t|t}$, in the bottom panels along with their 95% confidence interval.

According to the trend estimates for UC(r,0), trend output is above actual output at the beginning of the 70ies; it peaks at 1973.3 and starts declining until it reaches a trough in 1974.2. During this decline we observe a positive and high cycle, as implied by the strong and negative correlation between the two. A similar behaviour is found around 1979 and 1991. On the other hand, the final trend estimates for UC(0,θ) are less "volatile" and less related to the cyclical component; we adopt this terminology since even in a doubly infinite sample the estimates of the components will be correlated - the correlation can be computed from the elements of the matrix $\boldsymbol{P}_{t|\infty}$.

The two models produce very different smoothed cycle estimates: the estimates of the AR parameters in the UC(0,θ) case imply a stationary root at the zero frequency with multiplicity 2, whereas those for UC(r,0) imply a short run cycle with a period of about three years. It is also remarkable the difference between the real time, $\tilde{\psi}_{t|t}$, and final estimates of the cycle, $\tilde{\psi}_{t|T}$, especially for the model with correlated components. When the cycle is estimated in real time (bottom left panel), UC(r,0) lends support to the notion that this component represents a minor source of variation; we recall that the real time estimates cycle $\tilde{\psi}_{t|t}$ arising from UC(r, 0) is coincident with the BN cycle extracted from the unrestricted ARIMA(2,1,2). The latter is characterised by a perfect negative correlation (persistence is greater than 1) and has a non invertible ARMA(2,1) representation; as matter of fact, the parameter values reported in Table 1 for EA imply a value for the ϑ^* coefficient in (5) that is equal to -1.02.

However, for UC(r,0) the cycle is estimated in real time with non zero mean square error and the picture changes radically as we proceed to construct the final estimates using also future observations. These contradict the assertion that the cycle has a small amplitude, as it ranges from about -2.4% to +4.0%, as a percentage of GDP. Moreover, the final estimates have a much reduced standard error as compared to the real time ones. In particular, the increase in the reliability of the cycle estimates using a doubly infinite sample is as large as 88%. This quantity is defined as the percentage reduction in the estimation error variance when we compare the real time estimates with the final ones and is computed as: $100[\bar{P}_{t|t}^{(\psi)} - P_{t|\infty}^{(\psi)}]/\bar{P}_{t|t}^{(\psi)}$, where, using results presented in Section 7, and in particular (18), $\bar{P}_{t|t}^{(\psi)}$ is the steady state estimation error variance of the real time cyclical component and $P_{t|\infty}^{(\psi)}$ is that of the corresponding final estimates, using a doubly infinite sample.

If the BN decomposition is estimated as a *model*, that is we set up a state space model consisting of equations (3)-(5), after processing a suitable small number of observations the real time and final estimates are fully coincident. On the other hand, the estimates arising from UC(r,0) are subject to large revisions as new observations become available: indeed, for the estimated variance ratio $\sigma_\eta^2/\sigma_\kappa^2$ and AR parameters, a negative r implies that the distribution of the weights for extraction of the cycle, based on a doubly infinite sample, are highly skewed towards the future.

It is remarkable that in the more extreme case, when $r = -1$, the cycle is estimated with zero mean square error using a doubly infinite sample, but the real time estimates are characterised by high uncertainty, and we get a 100% increase in reliability from processing future observations. The model has a single source of disturbances, but it implies a non invertible ARIMA(2,1,2) representation, and thus the latter is not an innovation, but can be written as a linear combination of the current and future values of Δy_t.

On the contrary, when $r = 1$, the single source of disturbances is an innovation in a strict sense; it can be checked that the ACGF identity $g_r(L) = g(L)$ admits the solution $r = 1$, $\sigma_\eta = \sigma\theta(1)/\phi(1)$, $\sigma_\kappa = \sigma[1 - \theta(1)/\phi(1)]$, which implies $\phi_2\theta(1) + \theta_2\phi(1) = 0$. Thus, the real time estimates have the same AR(2) representation as the true component ($\vartheta^* = 0$); the process generating them is coincident with the maintained model for the unobserved component, so that current and past (i.e. real time) information is all we need to form this estimate.

In conclusion, if we accept that trend and cyclical disturbances are negatively correlated, then we must be willing to accept also that essential information for assessing the cyclical pattern lies in future observations and thus that our signals are prone to high revisions.

Figure 4 presents the estimated components of Italian GDP. The overall comments are unchanged except for the fact that the cycles extracted by UC(r,0) and UC(0,θ) have in this case about the same periodicity. We also present two alternative characterisations of macroeconomic fluctuations arising from the cyclical growth model, which is observationally equivalent to UC(r,0) and to the unrestricted ARIMA model. The first is the deviation cycle extracted under the hysteresis hypothesis (see model (10)); the latter modifies the trend permanently since it is integrated in the trend with a weight of 0.59. The second is the smoothed cycle in Δy_t based on (9). We can only resort to our prior to attach a preference to these alternative representations.

9 The Role of Seasonal Adjustment

The analysis of macroeconomic fluctuations usually relies on quarterly seasonally adjusted series. This raises the obvious issue as to whether seasonal adjustment can be considered as a neutral operation, in the sense that it does not alter the main stylised facts. The presence of a correlation between trend and cycle disturbances is one of those facts, given the relevance that the literature attaches to it.

To investigate this issue we perform a very simple Monte Carlo experiment, by which 1000 series of length $T = 140$ are generated according to $y_t = \mu_t + \psi_t + \gamma_t$, with independent trends and cycles, represented as UC(0,0) in (1); γ_t is a quarterly seasonal component, with trigonometric representation: $\gamma_t = \gamma_{1t} + \gamma_{2t}$, resulting from the sum of an annual non stationary cycle $(1 + L^2)\gamma_{1t} = \varpi_{1t}$ and a biannual one, $(1 + L)\gamma_{2t} = \varpi_{2t}$, with $\varpi_{1t} \sim \text{NID}(0, \sigma^2_\omega)$ and $\varpi_{2t} \sim \text{NID}(0, 0.5\sigma^2_\omega)$, independently of each other and of η_t and κ_t.

The cycle autoregressive parameters are written as $\phi_1 = -2\cos\lambda_c, \phi_2 = \rho^2$, where $\rho = 0.9$ and λ_c can take the two values $2\pi/12$ and $2\pi/32$ corresponding to a period of 3 (12 quarters) and 8 years (32 quarters), respectively. The trend-seasonal signal ratio is always kept at $\sigma^2_\eta/\sigma^2_\omega = 20$, whereas for $\sigma^2_\eta/\sigma^2_\kappa$ we consider three values, Low: $\sigma^2_\eta/\sigma^2_\kappa = 1/3$; Medium: $\sigma^2_\eta/\sigma^2_\kappa = 3$; High: $\sigma^2_\eta/\sigma^2_\kappa = 30$. The combination of these values with the two cycle periods gives 6 data generating processes in total.

For each simulation we fit the true model model and construct a seasonally adjusted (SA) series by removing from the simulated series the smoothed estimates of the seasonal component; the UC(r,0) is the fitted to the series. In the presentation of the results we label this experiment as SA-UC(r,0). Moreover, to characterise the small sample distribution of the correlation coefficient when the true value is $r = 0$, we estimate model a trend plus cycle plus seasonal model with correlated trend and cycle disturbances, that is UC(r, 0) plus an orthogonal seasonal component. We shall refer to this experiment with TCS(r, 0).

Figure 5 plots the distribution of the estimated r for SA-UC(r,0) and TCS(r,0) in the six cases. The histograms clearly point out that seasonal adjustment biases the estimates of the correlation coefficient, increasing the the evidence for a negative correlation. In general, the problem is lessened as we move away from the fundamental seasonal frequency (a yearly cycle), as the histograms for the 32 quarters cycle suggest.

Also, the panels in the second and the fourth columns highlight that the small sample distribution of r estimated on the unadjusted data is highly nonstandard, suffering from the same pile-up problem at ± 1 that was observed for the bootstrap distribution in Section 5. Experimentation suggests that we need a much larger sample size to have r distributed symmetrically around its true zero value.

10 The Role of Temporal Aggregation

Violation of the conditions under which a series admits an orthogonal trend-cycle decomposition may well be the consequence of temporally aggregating a flow variable. On the other hand, the only way in which systematic can affect the decomposability of the model into orthogonal components is via the small sample properties of the parameter estimates.

In this section we assume that observations are available on an aggregate series $Y_n, n = 1, 2, \dots, N$, that is obtained either by systematically sampling a stock variable or by aggregating a flow. Let s denote the aggregation period; if we denote the disaggregated series by $y_t, t = 1, 2, \dots, Ns$, in the former case $Y_n = y_{sn}$; the latter can be viewed as a systematic sample of $Y_n = \sum_{j=0}^{s-1} y_{sn-j} = S(L)y_t$, $S(L) = 1 + L + \cdots + L^{s-1}$, where the sample is taken at times $t = ns$;

Let us consider systematic sampling first: since $Y_n - Y_{n-1}$ is a systematic sample of $\Delta_s y_t$, its SGF is related to that of Δy_t, $g_{\Delta y}(\lambda)$, via the expression (Harvey, sec. 6.3.5):

$$g_{SS}(\lambda) = \frac{1}{s}\sum_{j=0}^{s-1} |S(e^{-\imath\omega_j})|^2 g_{\Delta y}(\omega_j), \tag{19}$$

where $\omega_j = s^{-1}(\lambda + 2\pi j)$, and

$$|S(e^{-\imath\omega_j})|^2 = S(e^{-\imath\omega_j})S(e^{\imath\omega_j}) = \begin{cases} \frac{1-\cos\omega}{1-\cos\omega_j}, & \omega_j \neq 0 \\ s, & \omega_j = 0 \end{cases}$$

is the power transfer function of the filter $S(L)$ evaluated at ω_j. Result (19) follows from application of the well known *folding* formula to the process $S(L)y_t$.

If the disaggregated series follows an UC(0,θ) model, so that the SGF of Δy_t is $g_{\Delta y}(\lambda) = \sigma^2_{\eta*} + 2(1-\cos\lambda)g_\psi(\lambda)$, where $g_\psi(\lambda)$ is the SGF of the cyclical component, then, using $\sum_{j=0}^{s-1} |S(e^{-\imath\omega_j})|^2 = s^2$, (19) specialises as

$$g_{SS}(\lambda) = s\left[\sigma^2_{\eta*} + 2(1-\cos\lambda)\frac{1}{s}\sum_{j=0}^{s-1} g_\psi(\omega_j)\right] \tag{20}$$

and the aggregated series can still be decomposed into orthogonal trend and cycle, with SGF $s^{-1}\sum_{j=0}^{s-1} g_\psi(\omega_j)$.

In the case of temporal aggregation, $Y_n - Y_{n-1}$ is a systematic sample of $\Delta_s S(L)y_t$ and thus the SGF of the aggregated series, $g_{TA}(\lambda)$, will be related to that of the disaggregated process as follows:

$$g_{TA}(\lambda) = \frac{1}{s}\sum_{j=0}^{s-1} |S(e^{-\imath\omega_j})|^4 g_{\Delta y}(\omega_j) \tag{21}$$

When the disaggregated series is an UC(0,θ) process, (21) becomes:

$$g_{TA}(\lambda) = s^{-1}\left\{ s^4\sigma^2_{\eta*} + (1-\cos\lambda)^2 \sum_{j=0}^{s-1}(1-\cos\omega_j)^{-2}\left[\sigma^2_\eta + 2(1-\cos\omega_j)g_\psi(\omega_j)\right]\right\} \tag{22}$$

Expression (22) has a complicated form and the decomposability into orthogonal components will arise only under very special conditions. The reduced form will be, in general, ARIMA(2,1,3) and will be decomposable into a RW trend plus ARMA(2,1) cycle plus irregular with correlated disturbances or into a RW trend plus ARMA(2,2) cycle with correlated components. Since temporal aggregation is a linear operation, another option is to decompose Y_n into an orthogonal IMA(1,1) trend and ARMA(2,2) cycle.

The ambiguities that aggregation of flow variables creates can be seen by working out the state space representation of Y_n (see Harvey, 1989). When the disaggregated model is (12), the state space model for the aggregate is (see Harvey, 1989):

$$\begin{array}{rcl} Y_n &=& \boldsymbol{z}'(\sum_{j=1}^{s}\boldsymbol{T}^j)\boldsymbol{\alpha}_{n-1} + \boldsymbol{z}'\left[\sum_{j=0}^{s-1}\sum_{i=0}^{j}\boldsymbol{T}^i\right]\boldsymbol{c} + u_t, \qquad n = 1,2,\dots,N, \\ \boldsymbol{\alpha}_n &=& \boldsymbol{T}^s\boldsymbol{\alpha}_{n-1} + (\sum_{j=0}^{s-1}\boldsymbol{T}^j)\boldsymbol{c} + \tilde{\boldsymbol{\epsilon}}_n, \end{array} \tag{23}$$

where

$$u_t = \boldsymbol{z}'\sum_{j=0}^{s-1}(\sum_{i=j}^{s-1}\boldsymbol{T}^{i-j})\boldsymbol{R}\boldsymbol{\epsilon}_{ns-j}, \quad \tilde{\boldsymbol{\epsilon}}_n = \sum_{j=0}^{s-1}\boldsymbol{T}^j\boldsymbol{R}\boldsymbol{\epsilon}_{ns-j}.$$

The representation (23) is already in the future state space form as can be seen on defining $\boldsymbol{\alpha}^*_n = \boldsymbol{\alpha}_{n-1}$. The new feature is the presence of the disturbance u_t in the measurement equation, that is correlated with the state disturbances. Interpreting u_t as correlated measurement noise is possible, but arbitrary. Other options, such as incorporating u_t into one of the components are arbitrary as well. Also, the BN decomposition will always estimate a RW trend, but the temporal aggregation of the trend component will give and IMA(1,1) process. These indeterminacies are resolved if the model is specified at the disaggregated frequency and estimated on the available series.

This leads us to the main point of this section. Suppose that the underlying model is UC(0,0) at the monthly level, but observations are available on the quarterly aggregate ($s = 3$). Can temporal aggregation explain the stylised fact that the spectral density of $Y_n - Y_{n-1}$ is not a minimum at the zero frequency? For this purpose we consider the Italian GDP and we fit (22) to the sample periodogram by maximum likelihood, as illustrated in Section 5.

The maximised likelihood is -140.38 and the parameter estimates are $\hat{\sigma}^2_{\eta*} = 0.0246$, $\hat{\sigma}^2_{\kappa*} = 0.0001$, $\hat{\phi}_1 = 1.94$, $\hat{\phi}_2 = -0.97$, implying a period of about 3 years (11 quarters), which amounts to the same period estimated by the quarterly UC(r,0). The model provides a good fit $Q(12) = 7.37$ normality 2.78, and yields a slightly greater likelihood than UC(r,0) (compare Table 1). Figure 6 shows how the model fits the raw periodogram; for comparison we report the parametric spectral density of UC(r,0). We notice that the spectral density estimate at zero and the implied persistence is the same, but the temporally aggregate UC(0,0) model (referred to as TA-UC(0,0)) has a sharper peak at the cyclical frequency. The estimated cycle, resulting from the aggregation of the monthly cycle, is plotted in the bottom panel and it is, roughly speaking, a compromise between those estimated by the quarterly UC(r,0) and by UC(0,θ).

The empirical findings in Rossana and Seater (1995) illustrate that temporally aggregated flows systematically show higher long run persistence with respect to the underlying disaggregated data. These results are not only a reflection of the small sample properties of the estimates, but also a theoretical consequence of temporal aggregation of flows. Persistence is the square root of the ratio $g(0)/\sigma^2$, where

$$\sigma^2 = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} g(\lambda)d\lambda\right)$$

is the prediction error variance (p.e.v.). Now, while it is straightforward to establish $g_{SS}(0) = sg_{\Delta y}(0)$ and $g_{TA}(0) = s^3 g_{\Delta y}(0)$, it proves difficult to derive analytical results on the effects of aggregation on persistence. However, we can easily prove that persistence of an aggregated flow is no smaller than that of a systematic sample. Applying the Cauchy-Schwartz inequality to the p.e.v. of an aggregated flow, it follows that $\sigma^2_{TA} \leq s^2\sigma^2_{SS}$, moreover, since $g_{TA}(0) = s^{-2}g_{SS}(0)$, we establish the result: $g_{TA}(0)/\sigma^2_{TA} \geq g_{SS}(0)/\sigma^2_{SS}$.

In conclusion, temporal aggregation of flow variables is non neutral with respect to the main stylised facts concerning macroeconomic fluctuations, such as persistence and correlated disturbances.

11 Multivariate Analysis

The previous analyses have been typically univariate. The issue that needs to be addressed at this stage is whether bringing in more information about the nature of economic fluctuations using related series can cast some light on correlated disturbances.

We set off reviewing some previous empirical results. Clark (1989) estimated a bivariate model of U.S. real output and unemployment grounded on the relationship between cyclical movements in output and unemployment known as Okun's law. The model for output is UC(r,0), and the unemployment rate is decomposed into a (driftless) random walk trend, unrelated to that in output, and a cyclical component that is a linear combination of the current and past value of the cycle in output. Clark estimated the correlation coefficient r to be equal to a nonsignificant -.12, with asymptotic 90% confidence interval (-0.4,0.3).

Jäger and Parkinson (1994) estimated bivariate UC models of real GDP and unemployment to examine the presence of hysteresis, according to which cyclical unemployment has an effect on the natural (trend) rate. They find that hysteresis effects are negligible in explaining the dynamics of U.S. unemployment, but are substantial for the Canadian, German and the U.K. unemployment rates.

Proietti, Musso and Westermann (2002) estimated a multivariate model made up of five time series equations for the Euro area Solow's residual, the labour force participation rate, the unemployment rate, capacity utilisation and the consumer price index, implementing the production function approach, augmented by a triangle model of inflation (see Gordon, 1997), to the measurement of potential output and the output gap for the Euro Area. They entertain hysteresis models and find mixed evidence for the unemployment series. They, however, prefer a specification featuring pseudo-integrated cycles that is at least as effective in explaining the persistence of the labour market variables.

We now proceed to a bivariate illustration concerning the Euro area GDP and consumer prices (p_t, logarithms). Within this framework the cycle in output takes the more specialised notion of an output gap, a measure of inflationary pressures, and the trend is the level of output that is consistent with stable inflation (potential output). The model is made up of the output equation, which is alternatively specified as UC(0,0) and UC(r,0), as given in (1), and the price equation is a structural version of Gordon's triangle model of inflation, specified as follows:

$$
\begin{aligned}
p_t &= \tau_t + \gamma_t + \textstyle\sum_k \delta_k x_{kt} \\
\\
\tau_t &= \tau_{t-1} + \pi^*_{t-1} + \eta_{\pi t} & \eta_{\pi t} \sim \text{NID}(0, \sigma^2_{\eta\pi}), \\
\pi^*_t &= \pi^*_{t-1} + \theta_\pi(L)\psi_t + \zeta_{\pi t} & \zeta_{\pi t} \sim \text{NID}(0, \sigma^2_{\zeta\pi}).
\end{aligned}
$$

where the regressors are commodity prices x_{kt} and the nominal effective exchange rate of the Euro, a level shift variable for 1974.1, and γ_t is a quarterly seasonal component, see Section 9. The only link between the prices and output equations is the presence of ψ_t as a determinant of underlying inflation, π^*_t, where $\theta_\pi(L) = \theta_{\pi 0} + \theta_{\pi 1}L$.

Table 2 reports selected estimation results concerning the model assuming uncorrelated disturbances (first column) and that with correlated ones (second column).

The estimated correlation coefficient is -0.26 and the likelihood ratio of the restriction $r = 0$ is not significant. What is more, unlike the univariate case, the estimates of the trend and cycle in output closely agree with those of the model with uncorrelated disturbances. These are displayed in Figure 7, along with estimates of underlying inflation, which is that part of observed inflation, devoid of seasonal fluctuations, related to the output gap, which is identified as the component π^*_t in the price equation. The autoregressive parameter estimates imply for both models a period of 10 years. It is also remarkable the reduction in the estimation error variance of the component ψ_t, compared to the univariate case (see Figure 3).

12 Concluding Remarks

The main conclusions of this paper is that the characterisation of macroeconomic fluctuations is by and large an open issue. On the one hand, models with correlated disturbances seem to improve the fit of UC decompositions to macroeconomic time series; this is true at least for the Italian GDP.

Table 2: Parameter estimates and diagnostics for bivariate models of quarterly euro area log GDP (y_t) and the logarithm of the consumer price index (p_t), 1970.1-2002.2. Standard errors in parenthesis

	UC(0,0)	UC(r,0)
σ^2_η ($\sigma^2_{\eta^*}$)	0.2231	0.2546
σ^2_κ ($\sigma^2_{\kappa^*}$)	0.0713	0.0987
ϕ_1	1.67	1.66
s.e	(0.05)	(0.05)
ϕ_2	-0.71	-0.71
s.e	(0.11)	(0.10)
r	0(r)	-0.26
s.e	-	(0.36)
$\sigma^2_{\eta\pi}$	0.0476	0.0468
$\sigma^2_{\zeta\pi}$	0.0000	0.0000
σ^2_ω	0.0000	0.0000
$\theta_{\pi 0}$	0.22	0.20
s.e	(0.04)	(0.05)
$\theta_{\pi 1}$	-0.20	-0.18
s.e	(0.04)	(0.04)
	Diagnostics and goodness of fit	
loglik	-133.28	-133.04
$Q(8)$ y_t	10.70	11.33
$Q(8)$ p_t	6.43	6.99
Normality y_t	9.18	9.51
Normality p_t	5.91	5.34

This finding is in part self-evident, since we entertain a more general model, and needs be interpreted with the following *caveat*: all our results are conditional on a particular ARIMA reduced form, which is itself an additional source of uncertainty in real life. For instance, Harvey and Jäger (1994) entertained an orthogonal trend-cycle decomposition to the U.S. real GDP series, allowing for a stochastic slope in the trend, so that the latter is an $I(2)$ process. Discriminating among UC models unconditionally, i.e. without assuming a particular reduced form, is a far more complex issue, due to the unavailability of a common estimable reduced form. On the other hand, several additional points were raised and illustrated, that mitigate this finding and are hereby summarised:

- Given the sample sizes typically available for macroeconomic time series the properties of the sampling distribution of the correlation parameter raise great concern. It was shown that asymptotic inferences do not provide a reliable guidance over them.

- Models with correlated components raise several interpretative issues, as under certain conditions they result observationally equivalent to models that provide different and equally plausible explanations of the nature of macroeconomic fluctuations. For the Italian GDP, the cyclical growth model and the hysteresis model provide exactly the same likelihood estimates.

- When we come to investigate the consequences of having highly and negatively correlated disturbances for signal extraction, it turns out that similarity with the BN decomposition does not carry over to the smoothed inferences, the estimated models implying that most information about the components is carried by future observations. Large revisions are thus to be expected.

- Seasonal adjustment and temporal aggregation can significantly affect the findings about correlated disturbances. In particular, the estimates of the correlation between trend and cycle disturbances are biased towards high and negative values.

- Univariate time series analysis cannot be demanded to solve such a controversial issue. Multivariate analysis can help. Our illustration, concerning a bivariate model of output and prices shows that the estimate of the correlation is substantially reduced and that the cycle in output can be estimated with increased reliability.

- The statistical literature has attached much significance to the restrictive nature of models with orthogonal components. However, when single source of errors and innovation representations are considered as a model, several inconsistencies arise. The emphasis on the exact nature of the resulting decompositions and on the absence of revision is misplaced and potentially misleading.

References

Anderson, B. D. O., and J. B. Moore (1979): *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ.

Beveridge, S., and C. R. Nelson (1981): "A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to the measurement of the 'business cycle'," *Journal of Monetary Economics*, 7, 151-174.

Blanchard, O., and D. Quah (1989): "The dynamic effects of aggregate demand and aggregate supply disturbances," *American Economic Review*, 79, 653-58.

Box G. E. P., D. A. Pierce, and P. Newbold (1987): "Estimating trend and growth rates in seasonal time series," *Journal of the American Statistical Association*, 82, 276-282.

Brewer, K. R. W. (1979): "Seasonal adjustment of ARIMA series," *Économie Appliquée*, 1, 7-22.

Brewer, K. R. W., P. J. Hagan, and P. Perazzelli (1975): "Seasonal adjustment using Box-Jenkins models," *Bulletin of the International Statistical Institute, Proceedings of the 40th Session*, 31, 130-136.

Burridge, P., and K. F. Wallis (1988): "Prediction theory for autoregressive-moving average processes," *Econometric Reviews*, 7(1), 65-95.

Casals, J., M. Jerez, and S. Sotoca (2002): "An exact multivariate model-based structural decomposition," *Journal of the American Statistical Association*, 97, 553-564.

Clark, P. K. (1987): "The cyclical component of U.S. economic activity," *The Quarterly Journal of Economics*, 102, 797-814.

Clark, P. K. (1989): "Trend reversion in real output and unemployment," *Journal of Econometrics*, 40, 15-32.

de Jong, P. (1989): "Smoothing and interpolation with the state space model," *Journal of the American Statistical Association*, 84, 1085-1088.

Doornik, J. A. (2001): *Ox. An Object-Oriented Matrix Programming Language*, Timberlake Consultants Press, London.

Doornik, J. A., and H. Hansen (1994): "An omnibus test for univariate multivariate normality," Discussion paper, Nuffield College, Oxford.

Durbin, J., and S. J. Koopman (2001): *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford, UK.

Fagan, G., J. Henry, and R. Mestre (2001): "An area-wide model (AWM) for the Euro area," *ECB Working Paper*, No. 42,

Gardner, E. S. (1985): "Exponential smoothing: the state of the art," *Journal of Forecasting*, 4, 1-28.

Godolphin, E. J. (1976): "Discussion on the paper by P.J. Harrison and C. F. Stevens 'Bayesian Forecasting'," *Journal of the Royal Statistical Society, Series B*, 38, 238-239.

Godolphin E. J., and J. M. Stone (1980): "On the structural representation for polynomial-projecting predictor models based on the Kalman filter," *Journal of the Royal Statistical Society, Series B*, 42, 35-45.

Gordon, R. J. (1997): "The time-varying NAIRU and its implications for economic policy," *Journal of Economic Perspectives*, 11 (2), 11-32.

Harvey, A. C. (1989): *Forecasting, Structural Time Series and the Kalman Filter*, Cambridge University Press, Cambridge, UK.

Harvey, A. C., and A. Jäger (1993): "Detrending, stylized facts and the business cycle," *Journal of Applied Econometrics*, 8, 231-247.

Harvey, A. C., and S. J. Koopman (2000): "Signal extraction and the formulation of unobserved component models," *Econometrics Journal*, 3, 84-107.

Harvey, A. C., and T. Trimbur (2002): "General model-based filters for extracting trends and cycles in economic time series," *Review of Economics and Statistics* (forthcoming).

Hyndman, R. J., A. B. Koehler, R. D. Snyder, and S. Grose (2002): "A state space framework for automatic forecasting using exponential smoothing methods," *International Journal of Forecasting*, 18, 439-454.

Jäger, A., and M. Parkinson (1994): "Some evidence on hysteresis in unemployment rates," *European Economic Review*, 38, 329-342.

Kuttner, K. N. (1994): "Estimating potential output as a latent variable," *Journal of Business and Economic Statistics*, 12, 361-368.

Morley, J. C., C. R. Nelson, and E. Zivot (2002): "Why are Beveridge-Nelson and unobserved-component decompositions of GDP so different?," *Review of Economics and Statistics*, forthcoming.

Koopman, S. J. (1997): "Exact initial Kalman filtering and smoothing for non-stationary time series models," *Journal of the American Statistical Association*, 92, 1630-1638.

Koopman S. J., N. Shepard, and J. A. Doornik (1999): "Statistical algorithms for models in state space using SsfPack 2.2," *Econometrics Journal*, 2, 113-166.

Lippi, M., and L. Reichlin (1992): "On persistence of shocks to economic variables. A common misconception," *Journal of Monetary Economics*, 29, 87-93.

Nerlove, M. L., D. M. Grether, and J. L. Carvalho (1995): *Analysis of Economic Time Series: A Synthesis*, Revised edition, Academic Press Inc., New York.

Ord, J. K., A. B. Koehler, and R. D. Snyder (1997): "Estimation and prediction for a class of Dynamic nonlinear statistical models," *Journal of the American Statistical Association*, 92, 1621-1629.

Piccolo, D. (1982): "A Comparison of some alternative decomposition methods for ARMA models," in *Time Series Analysis: Theory and Practice I* , O.D. Anderson (Ed.), Norh-Holland, p. 565-582.

Proietti, T. (1995): "The Beveridge-Nelson decomposition: properties and extensions," *Journal of the Italian Statistical Society*, Vol. 4. n. 1.

Proietti, T. (2002a): "Forecasting With Structural Time Series Models," in Clements, M.P. and D. F. Hendry (eds.), *A Companion to Economic Forecasting*, Blackwell Publishers, Oxford, 105-132.

Proietti, T. (2002b): "Some reflections on trend cycle decompositions with correlated disturbances," *EUI Working paper.*

Proietti T., and A. C. Harvey (2000): "The Beveridge-Nelson smoother," *Economics Letters*, 67, 139-146.

Proietti T., A. Musso, and T. Westermann (2002): "Estimating potential output and the output gap for the Euro area: a model-based production function approach," *EUI Working paper* ECO/9/2002.

Snyder, R. D. (1985): "Recursive estimation of dynamic linear models," *Journal of the Royal Statistical Society, Series B*, 47, 272-276.

Stoffer, D. S., and K. D. Wall (1991): "Bootstrapping state space models: gaussian maximum likelihood estimation and the Kalman filter," *Journal of the American Statistical Association*, 86, 1024-1033.

Watson, M. W. (1986): "Univariate detrending methods with stochastic trends," *Journal of Monetary Economics*, 18, 49-75.

Whittle, P. (1983): *Prediction and Regulation by Linear Least Squares Methods*, second edition, Basil Blackwell, Oxford.

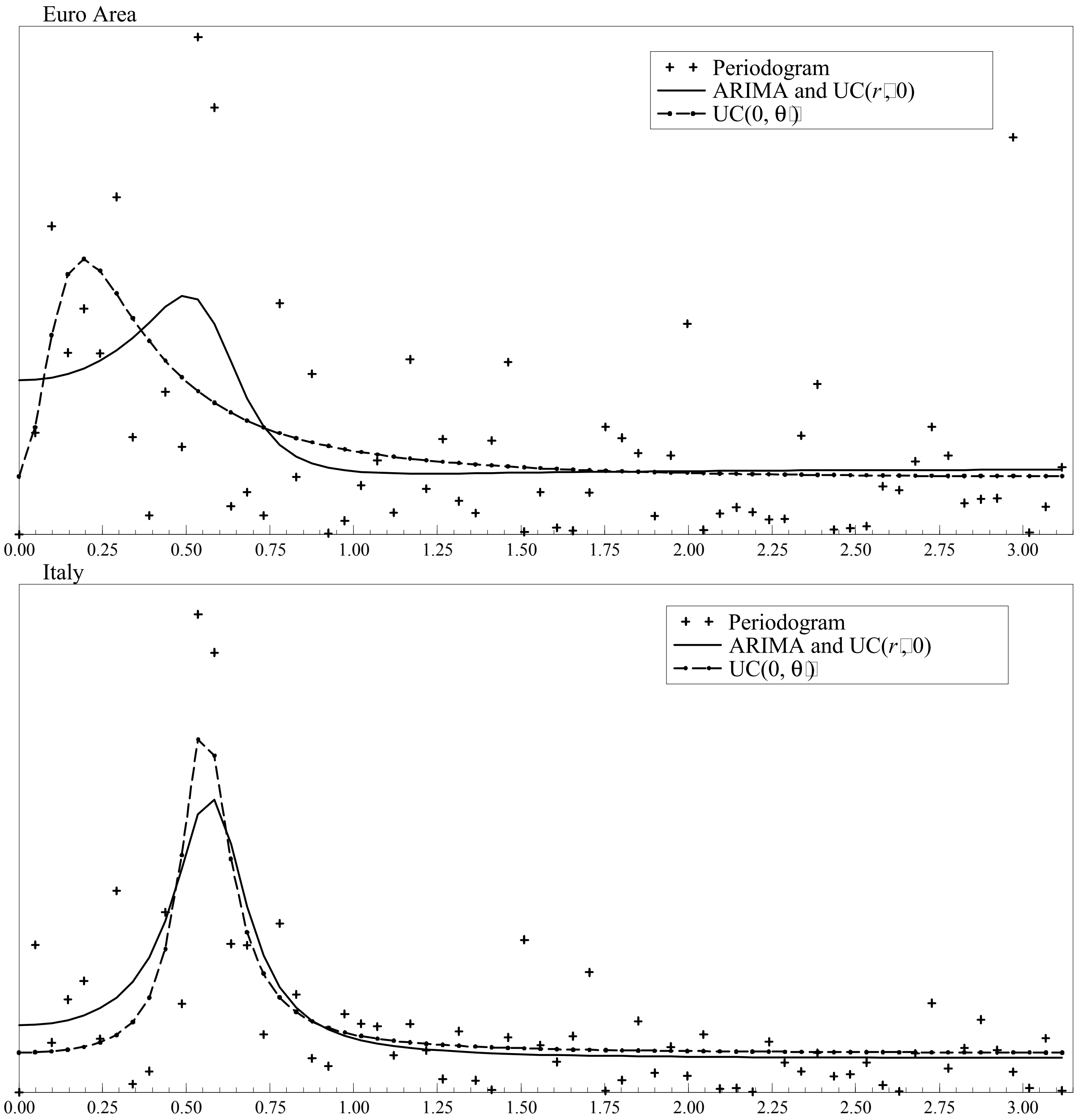

Figure 1: Euro Area and Italian GDP, 1970.1-2002.2. Periodogram, $I(\lambda_j)$, and parametric spectral densities of Δy_t, $g_m(\lambda_j)/(2\pi)$, estimated by the ARIMA(2,1,2) model, the UC(r,0), UC(0,θ)

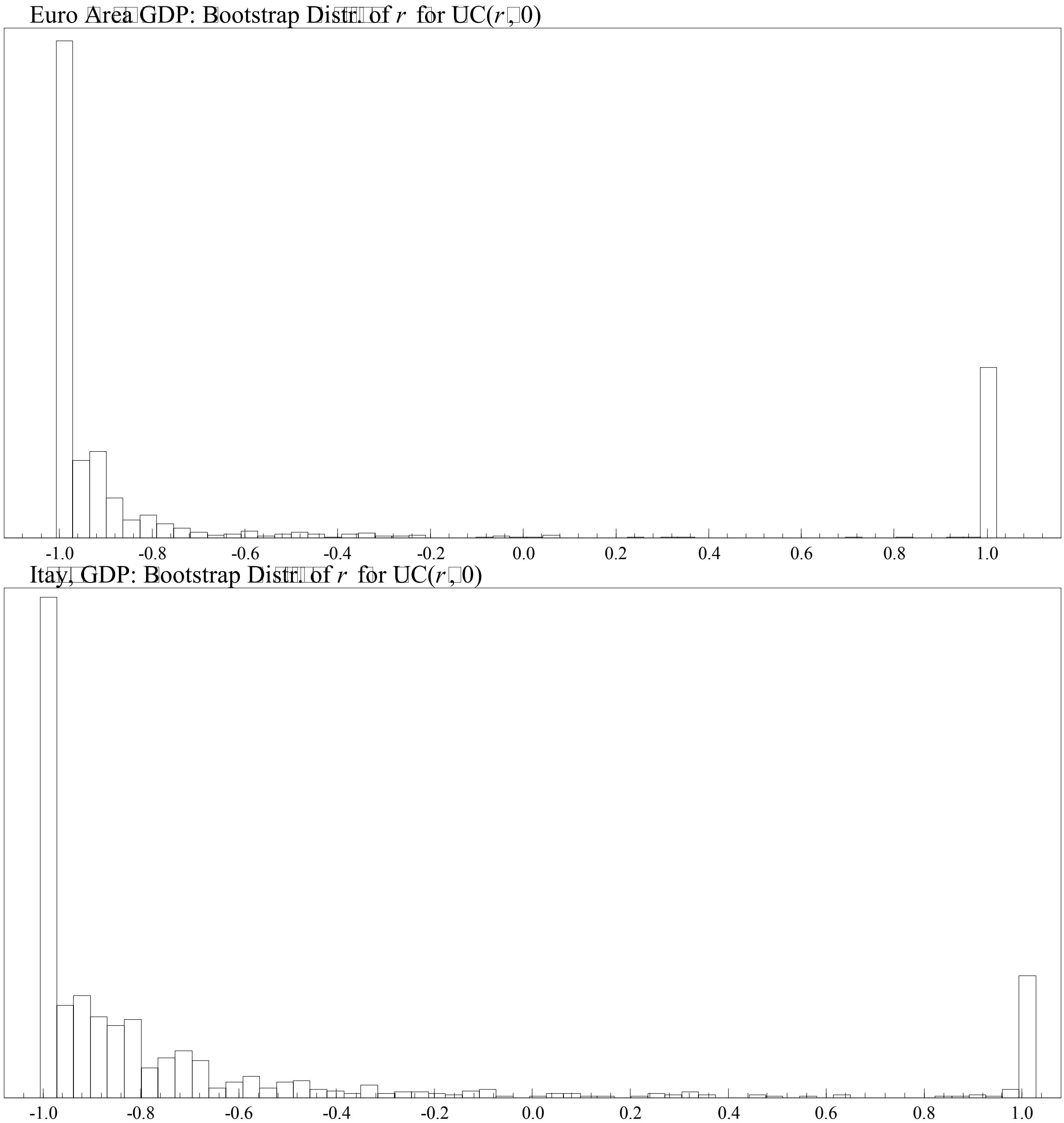

Figure 2: Euro Area and Italian GDP, 1970.1-2002.2. Distribution of the correlation parameter r in 1000 bootstrap samples

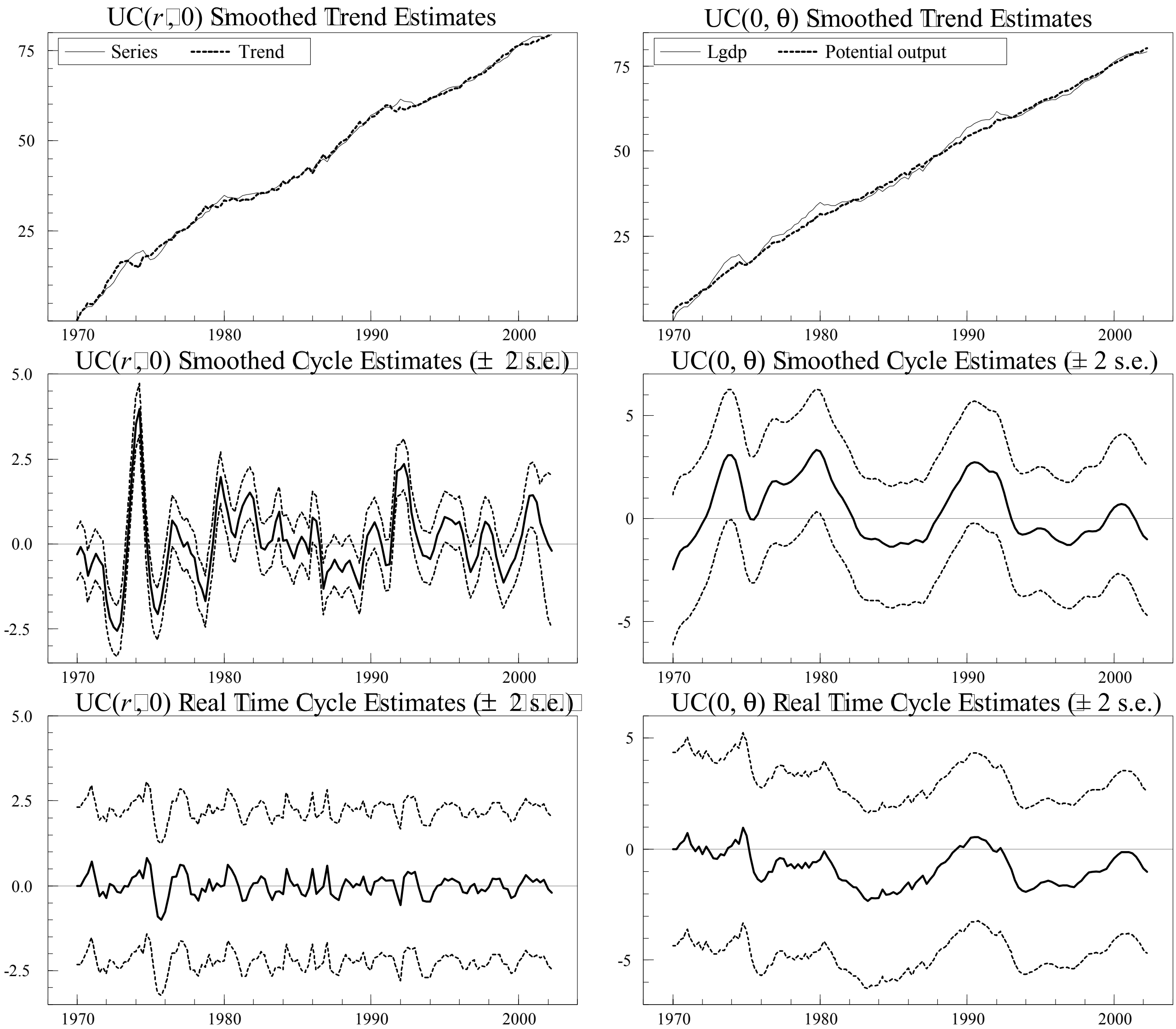

Figure 3: Euro Area GDP, 1970.1-2001.2. Smoothed estimates of trend, smoothed and real time estimates of the cycle arising from the UC(r, 0) model (left panels), UC(0, θ) (right panels)

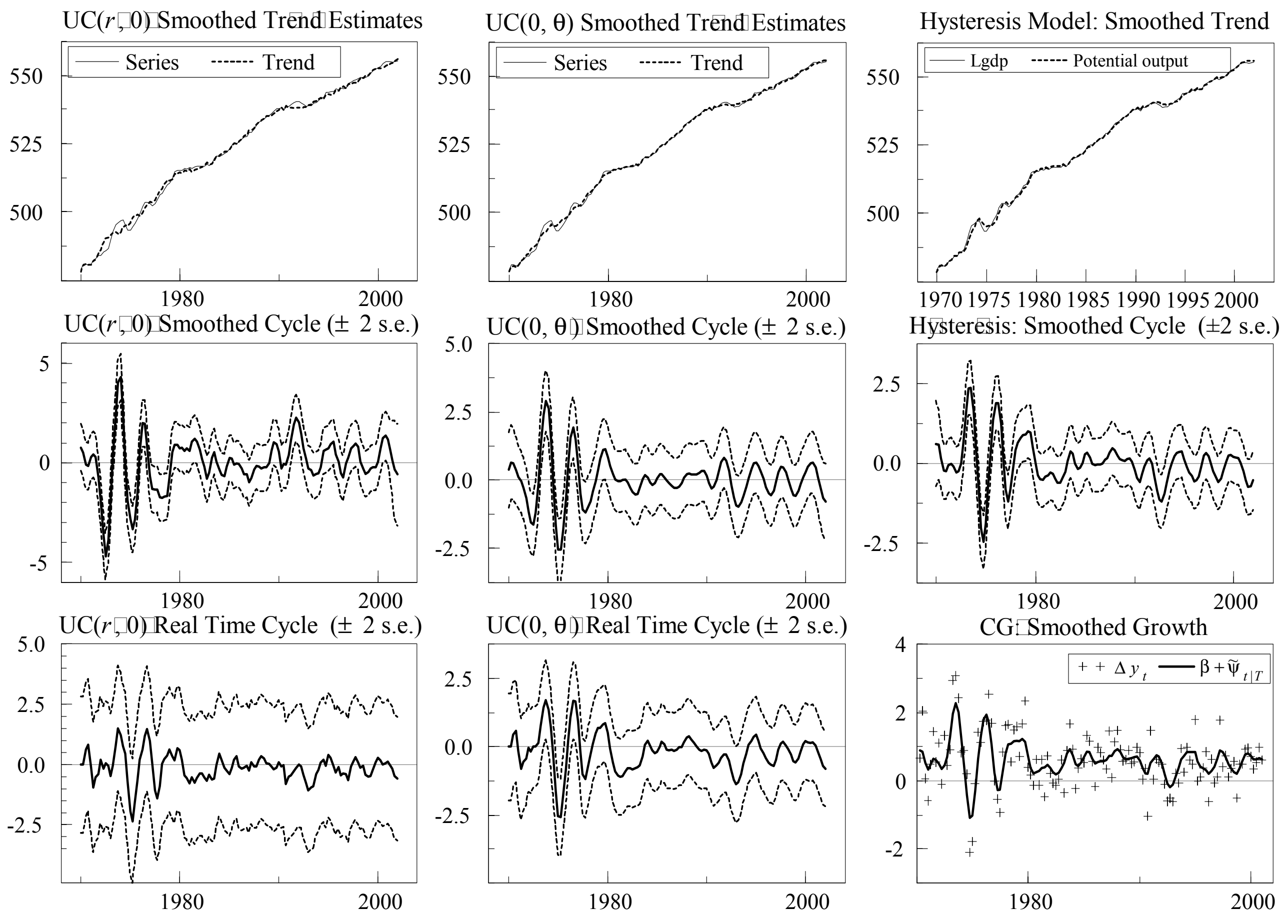

Figure 4: Italian GDP, 1970.1-2001.2. Smoothed estimates of trend, smoothed and real time estimates of the cycle arising from the UC(r,0) model (left panels), UC(0,θ) (centre panels). For the CG-hysteresis model we present the smoothed estimates of the trend (top right panel) and the cycle (middle right panel) and the smoothed estimates of underlying growth (CG model), $\beta + \tilde{\psi}_{t|T}$ (bottom right panel)

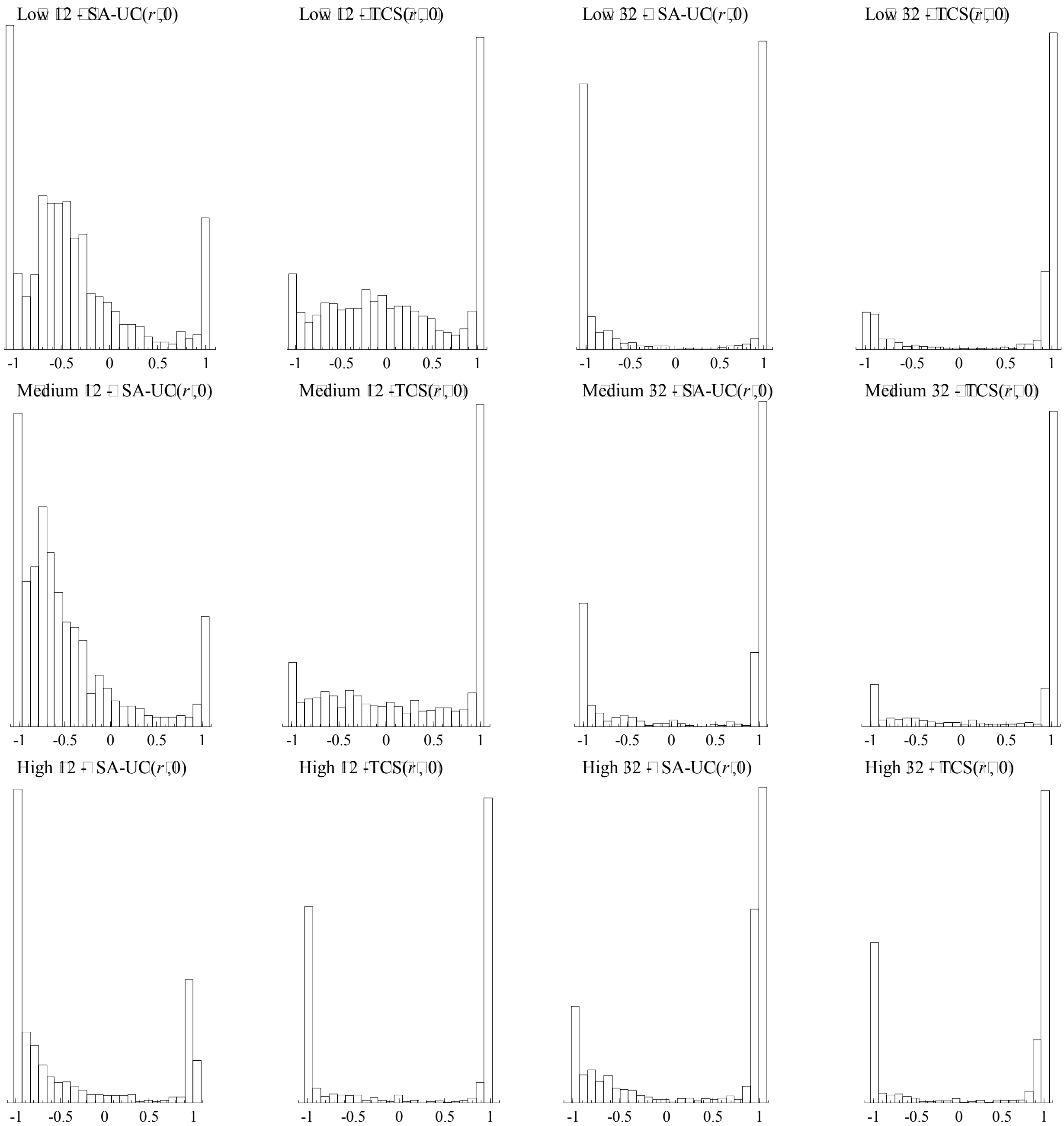

Figure 5: Distribution of the correlation coefficient, r, for the UC(r,0) model estimated on seasonally adjusted data and on the raw simulated data by fitting TCS(r,0). 1000 quarterly series of length $T = 140$ are generated according to orthogonal trend plus cycle plus seasonal models with low, medium and high signal ratios, and cycle periods equal to 12 and 32 quarters

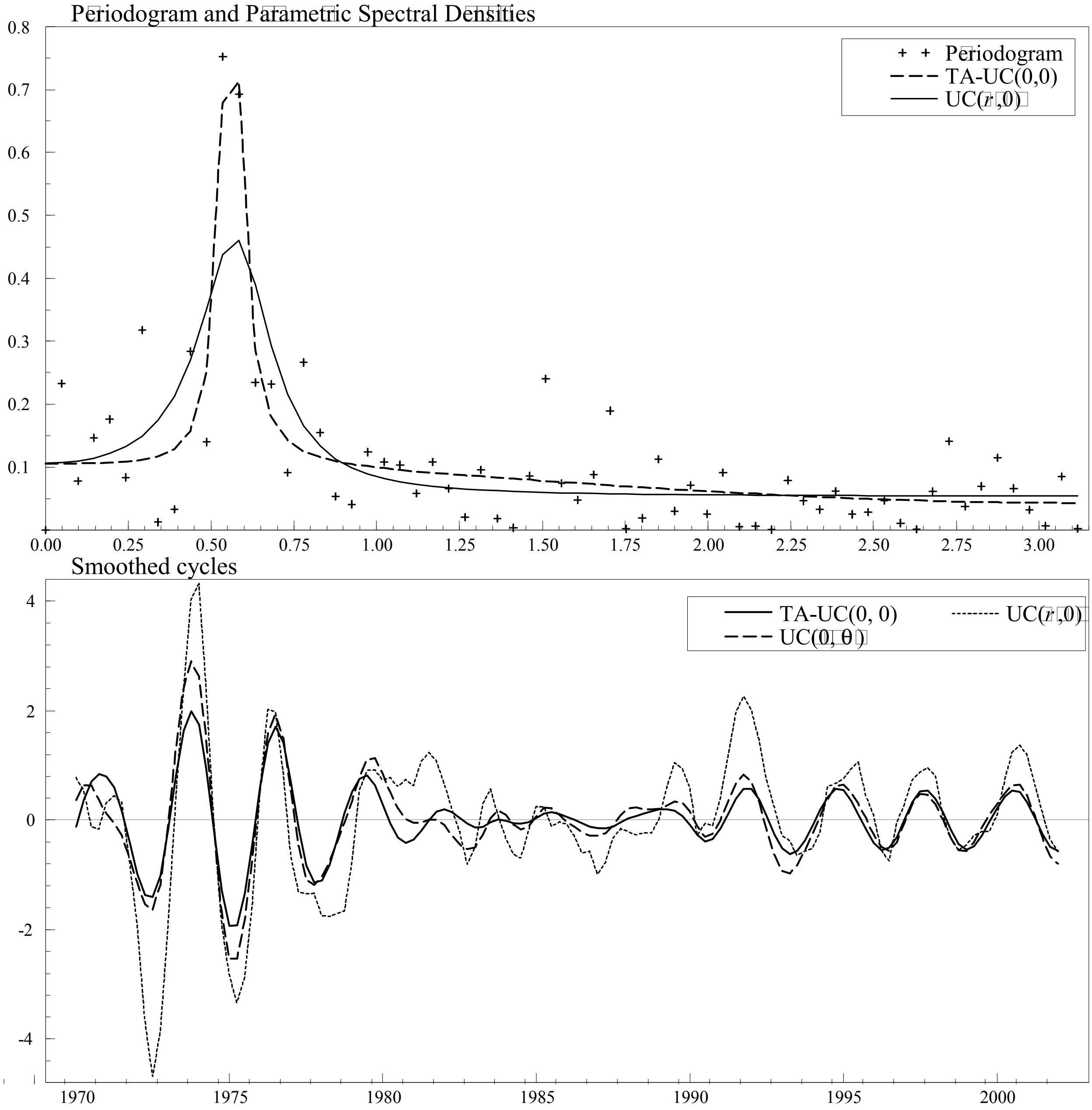

Figure 6: Italian GDP, 1970.1-2002.2. Periodogram, $I(\lambda_j)$, and parametric spectral densities of Δy_t, $g_m(\lambda_j)/(2\pi)$, estimated by the temporally aggregated TA-UC(0,0) model and UC(r,0)

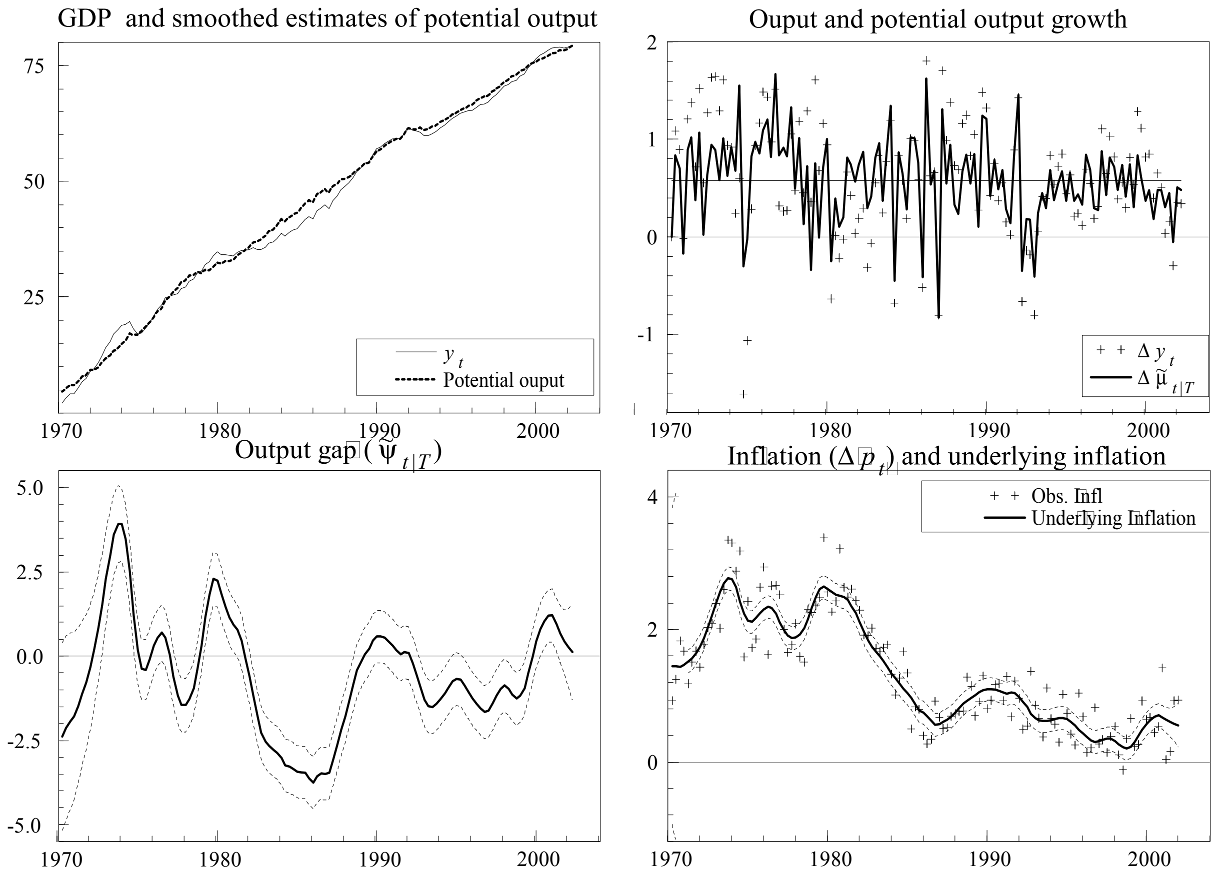

Figure 7: Euro Area GDP and consumer prices (logarithms). Estimates of potential output ($\tilde{\mu}_{t|T}$), potential output growth ($\Delta\tilde{\mu}_{t|T}$), the output gap ($\tilde{\psi}_{t|T}$), and underlying inflation, ($\tilde{\pi}^*_{t|T}$), with 95% confidence intervals, resulting from the bivariate model with correlated disturbances of Section 11