Exploring US Business Cycles with Bivariate Loops using Penalized Spline Regression

Göran Kauermann Timo Teuber
Peter Flaschel
Bielefeld University

September 19, 2006

Abstract

The phrase business cycle is usually used for short term fluctuations in macroeconomic time series. In this paper we focus on the estimation of business cycles in a bivariate manner by fitting two series simultaneously. The underlying model is thereby nonparametric in that no functional form is prespecified but smoothness of the functions are assumed. The functions are then estimated using penalized spline estimation. The bivariate approach will allow to compare business cycles, check and compare phase lengths and visualize this in forms of loops in a bivariate way. Morevover, the focus is on separation of long and short phase fluctuation, where only the latter is the classical business cycle while the first is better known as Friedman or Goodwin cycle, respectively. Again, we use nonparametric models and fit the functional shape with P-splines. For the separation of long and short phase components we employ an Akaike criterion.

JEL CLASSIFICATION SYSTEM FOR JOURNAL ARTICLES:
C32, **C14**, **E32**.

KEYWORDS: penalized spline regression, bivariate nonparametric regression, business cycle, Goodwin and Friedman cycle

1 Introduction

Econometrics as well as statistics have been dominated by parametric models over decades. The limitation has fallen down with smoothing techniques now allowing for a sophisticated framework to fit models functionally. Smoothing techniques with local approaches (see e.g. Fan and Gijbels 1996) as well as spline smoothing (see Eubank 1989 or Wahba 1990) have indeed achieved recognizable standard and are numerically available in modern statistics packages like **R** (www.r-project.org) or S-Plus. A general introduction is found in Hastie and Tibshirani (1990), a milestone in terms of smooth nonparametric modeling. In recent years, penalized spline fitting as smoothing technique has obtained more and more attention. Originally introduced by O'Sullivan (1986) it was Eilers and Marx (1996) who made the procedure popular under the phrase P-spline smoothing. The book by Ruppert, Wand, and Carroll (2003) demonstrates the flexibility as well as the numerical simplicity of the approach. An interesting feature of the P-spline idea is thereby, that P-spline smoothing can be linked to Linear Mixed Models so that both, fitting as well as smoothing parameter selection can be carried out by Mixed Models technology, see Wand (2003) or Kauermann (2004).

Even though nonparametric techniques are well developed and established in classical regression type models, their use for time dependent data is less explored. Fan and Yao (2003) summarize a number of nonparametric models in a time series framework. More recent work is found in Ruppert (2004) or the edited volume by Akritas and Politis (2003), chapters 8 and 11. A concise general overview was given in Härdle, Lütkepohl, and Chen (1997). Nonetheless, the field of nonparametric estimation for time series data has shown distinctly less activity in the last years compared to the ubiquitous discussion in a classical regression context. In this paper we make use of functional data analyses using penalized splines to structure and analyze business cycles and underlying long phase cycles of bivariate macroeconomic time series. The major problem one is facing in data collected over time is that residuals ε_t might be serially correlated. And even though serial correlation is not a particular problem for the pure nonparametric fit of $g(t)$, it is a problem for selecting the smoothing parameter in a data driven way, (see Opsomer, Wang, and Yang 2001). For P-spline smoothing the problem is less apparent as has been shown in Krivobokova and Kauermann (2004). Nonetheless, the decomposition of trend and serial correlation is a difficult problem and a uniform scientific opinion can not be achieved.

We investigate the relation of the employment rate y_{t1} and inflation rate y_{t2}. From a pure statistical point of view we are thereby not interested in a separate analysis and fitting of the two business cycle functions for y_{t1} and y_{t2}, respectively. We assume that $\mathbf{y}_t = (y_{1t}, y_{2t})$ follows a business fluctuation $\mathbf{g}(t) = (g_1(t), g_2(t))$, say, and pursue a joint estimation of $\mathbf{g}(t)$. The objective is to visualize $\mathbf{g}(t)$ as trajectory over t in the phase space of y_{t1} and y_{t2}. The long phase loops and the business fluctuations around them obtained from this trajectory allow for an economic interpretation. In particular the direction of rotation of the trajectory of $g_1(t)$, $g_2(t)$ gives information which cycle is running ahead of the other and clearly, a rotation change indicates, that the phase length for the two cycles $g_1(t)$ and $g_2(t)$ differ. We will use polar coordinates which allows to assess if cycle length differ and whether they have a dynamic behavior over time. Practically this is done by smoothly estimating rotation angle and radius of the function $\mathbf{g}(t)$. We will decompose the series into a long phase cycle and a short phase cycle of business cycle frequency, where not only the latter is of interest in this paper, though our angle and radius measures are characteristics of the observed business solely. Both, "short" and "long" phase cycles are estimated in a coherent framework using Penalized Splines. The modeling exercise has relations to classical predator-prey models originally proposed by Lotka (1925) and Volterra (1926) and applied to economics by Goodwin (1967). There loops or cycles, respectively, are trajectories of differential equations which are itself parametric. Our approach does not follow the differential equation relation but uses non-parametric routines to estimate smooth loops. The smoothing idea itself shows similarities to estimating principle curves (see e.g. Hastie and Stützle 1989, or Einbeck, Tutz, and Evers 2005, and references given there). Unlike the data situation in principle curve estimation, however, we have data collected over time which is in fact the core component of our model.

As data we investigate bivariate macroeconomic time series. In more detail we are interested in the short term fluctuations in the time series, which are well known as "business cycles". The most popular definition of business cycles is given by Burns and Mitchell (1946) which name them as "a type of fluctuations in the aggregate economic activity". Although the work was criticized by Koopmans (1947) the definition and nowadays the approaches are mainly accepted. Long and Plosser (1983) agree to this definition and stated that the name "business cycle refers to the joint time-series behavior of a wide range of economic variables". Instead of observing these fluctuations directly in the time plot of the time series one has to derive the "detrended" time series. Lucas (1977) redefines the business cycles as "the deviations of the Gross National Product from a trend", which can differ from an exponential growth rate over the time. Kydland and Prescott (1990) propose "a curve which students of business cycles and growth would draw in" and suggest to use the Hodrick and Prescott (1981) Filter. Contrary, Stock and Watson (1999) prefer the bandpass filter (Baxter and King 1999). From a statistical point of view both filter have a lack, namely the choice of the "tuning parameter". Hodrick and Prescott (1981) and Kydland and Prescott (1990) use a fixed smoothing parameter $\lambda = 1600$, which was subjectively proposed for one time series (real output) and it seems questionable to use this smoothing parameter blindly for all other time series. Baxter and King (1999) and Stock and Watson (1999) choose the tuning parameter pair $p = 6$ and $q = 32$ because they pointed out that a cycle should last at least 6 and at most 32 quarters long. Although the choices of these parameters are (subjectively) reasoned we again warn to follow this suggestions blindly. A statistical approach which selects the tuning parameter data driven is suggested in the paper.

The equilibrium models of Kydland and Prescott (1982), Long and Plosser (1983) and Backus, Kehoe, and Kydland (1992) build up a microfounded economy in which the artificial time series are very alike the empirical fluctuations measured by statistical moments, i.e. the covariance and autocorrelation. While the Kydland and Prescott (1982) model is based on the idea that the production of the good takes more than one period, Long and Plosser (1983) extend the one-good-model to an economy in which different goods are produced to be consumed and/or to produce new goods. Backus, Kehoe, and Kydland (1992) opens the economy in which labor is immobile but goods can be transferred between the countries to reduce the risk of the agents. Although these models differ slightly they are all capable to reproduce more or less the observed fluctuations of macroeconomic time series, e.g. real output, consumption, investment, trade balance, capital stock and worked hours.

Given the theoretical background of the equilibrium models more and more statistical approaches have been presented in the literature to capture the behavior of the business cycles. Stock and Watson (1999) compare seventy macroeconomic time series with the real output. Beside a graphical representation they capture the relationship by the autocorrelation statistics and the Granger causalities. Hamilton (1989) and Hamilton (2005) use a discrete Markov switching model to explain some time series. Sinclair (2006) uses this work to extend the model to explain the observed asymmetries in the business cycles. The work of Stock (1987) distinguish between the observed (linear) calendar time and a (nonlinear) economic time. These examples justifies a nonlinear approach although most of the works seem to fail to extend the models to a multivariate case.

The scientific contributions of this paper are twofold. First, from a methodological point of view, we propose a new econometric method for decomposing bivariate time series, and their representation as two-dimensional phase plots, into two cycles of significantly different length. Secondly, from an economic viewpoint, we provide the estimation and visualization of such interacting cycles of business cycle frequency and of long wave frequency for inflation and income distribution dynamics in its dependence on the rate of employment, respectively. These estimations can be viewed as laying the ground for a new investigation of time series with no secular trend by decomposing them into cyclical ”trends’ and ordinary business cycles simultaneously. In the case of the US economy after World War II we obtain approximately six business cycles that are superimposed on one long cyclical fluctuations in employment, inflation and income distribution. From a methodological point of view we here use smooth estimation of cycles and long loops, a field which is only rudimentary explored in statistics (see also Fisher 1995).

The paper is organized as follows. In section 2 we present the estimation routine with its theoretical background. The link to Generalized Linear Mixed Models is derived which is used for smoothing parameter selection. Section 3 discusses simulations as well estimation for our two bivariate data examples, the inflation and the income distribution cycles. Section 4 concludes.

2 Estimating bivariate Business Cycles

2.1 B-spline Estimation

Assume we observe data points (y_{t1}, y_{t2}) in pairs with t as index referring to the time point. We assume that the data are noisy observations of a smooth two dimensional function $\mathbf{g}(t) = (g_1(t), g_2(t))^T$, where $\mathbf{g}$ is smooth in the following sense. The trajectory $\mathbf{g}(t)$ follows loops or circles around the origin, and both, the velocity as well as the radius have no rapid changes. In particular this means, that locally and ignoring the implicit role of t, $g_1(\cdot)$ is a smooth function of $g_2(\cdot)$ and vice versa, respectively. More precisely we formulate $\mathbf{g}(t)$ in the polar coordinate functions

$$\begin{aligned} \text{radius:} \quad \rho(t) &= \sqrt{g_1(t)^2 + g_2(t)^2} && (2.1) \\ \text{angle:} \quad \phi(t) &= \arctan\left(\frac{g_2(t)}{g_1(t)}\right) \\ &+ 1_{\{g_1(t)<0\}}\left[1_{\{g_2(t)>0\}} - 1_{\{g_2(t)<0\}}\right]\pi && (2.2) \end{aligned}$$

which are smooth functions in t, where $1_{\{.\}}$ is the indicator function. Clearly, smoothness of $\rho(t)$ refers to smooth changes of the radius while smoothness of $\phi(t)$ means circular smoothness with jumps at -2π. Retransformation allows to write $\mathbf{g}(t)$ as

$$\mathbf{g}(t) = \Big(\rho(t)\cos\phi(t) \quad \rho(t)\sin\phi(t)\Big)^T \tag{2.3}$$

We assume now that $\mathbf{y}_t = (y_{t1}, y_{t2})^T$ are noisy observations of $\mathbf{g}(t)$, that is

$$\mathbf{y}_t = \mathbf{g}(t) + \boldsymbol{\epsilon}_t \tag{2.4}$$

with $\boldsymbol{\epsilon}_t = (\epsilon_{t1},\, \epsilon_{t2})^T$ as residuals. For simplicity, and to make the machinery of estimation running, we first assume that $\boldsymbol{\epsilon}_t$ are independent over time, but it seems necessary to allow for correlation between ϵ_{t1} and ϵ_{t2}. With normality assumed we denote this as $\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$ with $\boldsymbol{\Sigma}_\epsilon$ as covariance matrix.

Functions $\rho(t)$ and $\phi(t)$ are estimated using a P-spline approach in the style of Eilers and Marx (1996) and Ruppert, Wand, and Carroll (2003). We therefore set

$$\rho(t) = \exp\left(\tilde{\rho}(t)\right) = \exp\left\{\mathbf{B}_\rho(t)\mathbf{b}_\rho\right\} \tag{2.5}$$

where $\mathbf{B}_\rho(t)$ is a spline basis built over the support of t. The $\exp\{\cdot\}$ link in (2.5) is used for technical reasons to ensure a positive radius and $\tilde{\rho}(t)$ is the linear combination of the splines, i.e. $\tilde{\rho}(t) = \mathbf{B}_\rho(t)\mathbf{b}_\rho$. The spline basis is chosen in a rich manner with knots for spline functions placed every 4-5 observed time points. A more theoretical investigation on how many spline functions should be chosen asymptotically is provided in Ruppert (2002). In principle, the choice of the basis functions in $\mathbf{B}_\rho(\cdot)$ is left to the user and any spline shape function could be used. For simplicity, both in terms of numerical behavior and notation, we work with B-spline bases of third order as introduced in de Boor (1978). The B-spline basis is built from piecewise polynomials connected at knots $\tau_0, \tau_1, \ldots, \tau_K$. In our example we use equidistant knots covering the support of t. A short sketch on how B-splines are built is given in the Appendix. Analogously to the radius we model the basis for changes of the angle $\phi(t)$. This is accommodated by setting

$$\phi(t) = \mathrm{mod}\left(\tilde{\phi}(t)\right) = \mathrm{mod}\Big(\mathbf{B}_\phi(t)\mathbf{b}_\phi\Big) \tag{2.6}$$

where $\mathrm{mod}(x) = 2\pi\left(\frac{x}{2\pi} - \left\lfloor \frac{x}{2\pi} \right\rfloor\right)$ and $\lfloor x \rfloor$ returns the smallest integer value of x. Again $\tilde{\phi}(t)$ is the linear combination of the splines. Note that $\mathrm{mod}(\cdot)$ is used for graphical reasons only and the discontinuity is not a technical problem. In fact, we have for instance $\sin(\phi(t)) = \sin(\tilde{\phi}(t))$. Spline basis $\mathbf{B}_\phi(t)$ in (2.6) can in principle be chosen differently from $\mathbf{B}_\rho(t)$, but to keep the procedure simple we choose $\mathbf{B}_\phi(t) = \mathbf{B}_\rho(t)$. Assuming normality for the residuals we achieve to the log likelihood

$$l(\mathbf{b}, \mathbf{\Sigma}_\epsilon) = -\frac{n}{2} \log|\mathbf{\Sigma}_\epsilon| - \frac{1}{2}\sum_{i=1}^{n} (\mathbf{y}_i - \mathbf{g}(t_i))^T \, \mathbf{\Sigma}_\epsilon^{-1} \, (\mathbf{y}_i - \mathbf{g}(t_i)) \qquad (2.7)$$

with $\mathbf{b} = (\mathbf{b}_\rho^T, \mathbf{b}_\phi^T)$ and $\mathbf{g}(\cdot)$ as defined in (2.3). Simple parameter maximization of (2.7) would provide unsatisfactory estimates since bases $\mathbf{B}_\rho(t)$ and $\mathbf{B}_\phi(t)$ were chosen high dimensional and the corresponding estimates would be jagged. We therefore pursue a penalized fit instead, by imposing a penalty on $\mathbf{b}_\rho$ and $\mathbf{b}_\phi$, respectively. This is achieved by maximizing the penalized likelihood

$$l_P(\mathbf{b}, \mathbf{\Sigma}_\epsilon; \lambda_b) = l(\mathbf{b}, \mathbf{\Sigma}_\epsilon) - \frac{1}{2}\lambda_\rho \mathbf{b}_\rho^T \mathbf{D}_\rho \mathbf{b}_\rho - \frac{1}{2}\lambda_\phi \mathbf{b}_\phi^T \mathbf{D}_\phi \mathbf{b}_\phi \qquad (2.8)$$

with $\lambda_b = (\lambda_\rho, \lambda_\phi)$ as penalty parameters and $\mathbf{D}_\rho$ and $\mathbf{D}_\phi$ as penalty matrices. As has been suggested in Eilers and Marx (1996) a smooth fit is achieved if spline coefficients of adjacent B-splines are of the same order. This is achieved by imposing a penalty on first or higher order differences of the elements on $\mathbf{b}_\rho$ and $\mathbf{b}_\phi$, respectively. In the simplest case we penalize $b_{\rho l} - b_{\rho l-1}$, which can be written in matrix form as $\mathbf{L}\mathbf{b}_\rho$ with $\mathbf{L}$ as $(p-1) \times p$ dimensional contrast matrix where p is the dimension of $\mathbf{b}_\rho$. Setting now $\mathbf{D}_\rho = \mathbf{L}_\rho^T \mathbf{L}_\rho$ leads to the penalty matrix in (2.8). The same applies to the construction of $\mathbf{D}_\phi$.
Statistical properties of the estimate as well as optimization with respect to the smoothing parameter λ_ρ and λ_ϕ are listed in Appendix A.

2.2 Business Fluctuations and Long Phase Loops

We have assumed so far that the center for cycles described by $\mathbf{g}(t)$ is zero which implies that the series y_{t1} and y_{t2} are stationary without any long phase cycle. Apparently this is a stringent assumption which will be weakened now to a more practical situation. To do so we replace model (2.4) by

$$\mathbf{y}_t = \mathbf{c}(t) + \mathbf{g}(t) + \varepsilon_t \tag{2.9}$$

where $\mathbf{c}(t) = (c_1(t), c_2(t))^T$ is the long phase cycle around which $\mathbf{g}(t)$ is oscillating. In (2.9) we have now decomposed the mean structure into long phase movement $\mathbf{c}(t)$ and shorter phase oscillation $\mathbf{g}(t)$. In time series analysis the decomposition of trends and seasonal effects is well established (see e.g. Brockwell and Davis 1987). Yet in our situation we have cyclical trends, see the next section, as well as business fluctuations in seasonally adjusted time series data. Unlike in classical time series, the phase length of the these cycles are unknown and the objective is to estimate these from the data. The canonical candidate for long phase cycle estimation is the Hodrick and Prescott (1997) filter. It leaves, however, the unsatisfactory requirement of choosing a penalty parameter λ with its recommended setting $\lambda = 1600$. From a statistical point of view fixing the smoothing parameter in advance is unsatisfactory and a data driven criterium seems preferable. We therefore pursue a smooth approach by fitting $\mathbf{c}(t)$ again using penalized spline fitting, that is we replace $\mathbf{c}(t)$ by

$$\mathbf{c}(t) = \begin{pmatrix} \mathbf{Z}_1(t)\mathbf{a}_1 \\ \mathbf{Z}_2(t)\mathbf{a}_2 \end{pmatrix} =: \mathbf{Z}(t)\mathbf{a}$$

where $\mathbf{Z}_l(t)$ are spline bases chosen complex enough to capture long phase cycles, $l = 1, 2$. Using a B spline basis for $\mathbf{Z}_l(t)$ with same knots like for the estimation of $\mathbf{g}(t)$ the spline coefficient $\mathbf{a}$ is now estimated in a penalized form with penalty $\mathbf{a}_l^T \mathbf{L}_a^T \mathbf{L}_a \mathbf{a}_l, l = 1, 2$, and $\mathbf{L}_a$ as difference matrix. In principle we could now formulate the penalty as a priori normality and fit the resulting structured mixed model. To keep the numerics simple and understandable, we proceed however with a hybrid two step procedure. This means we first estimate the long phase cycle $\mathbf{c}(t)$ and then fit the business cycle structure $\mathbf{g}(t)$ to the residuals $\tilde{\mathbf{y}}_t = \mathbf{y}_t - \mathbf{c}(t)$. This hybrid approach appears justifiable since our objective is the estimation of the shorter phase structure $\mathbf{g}(t)$ in its dependence on the longer cycle. We therefore fit $\mathbf{c}_l(t)$ componentwise with given penalty parameters $\lambda_a = (\lambda_1, \lambda_2)$, say. That is

$$\mathbf{c}_l(t) = \mathbf{Z}_l(t) \left(\mathbf{Z}_l^T \mathbf{Z}_l + \mathbf{D}_a(\lambda_l) \right)^{-1} \mathbf{Z}_l^T \mathbf{Y}_l = \mathbf{S}_l(\lambda_l) \mathbf{Y}_l \qquad , l = 1, 2$$

with $\mathbf{Y}_l = (y_{1l}, \ldots, y_{nl})^T$ and $\mathbf{Z}_l$ as matrix built from $z_l(t_i)$, $i = 1, \ldots, n$ and $\mathbf{S}_l(\lambda_l)$ as smoothing matrix. In particular, for the long phase cycle we ignore any possible correlation among the components of $\mathbf{y}$. The resulting residuals $\tilde{\mathbf{y}}_t$ are assumed to be distributed according to (A.5) with $\mathbf{y}_t$ replaced by $\tilde{\mathbf{y}}_t$. In a second step estimate $\hat{\mathbf{g}}(t)$ is obtained based on the observations corrected by the long phase cycle. It remains to select appropriate penalty parameters for λ_a. We propose an Akaike based criterion here using a grid search for λ_a. This means we intend to find the minimum value for

$$\text{AIC}(\lambda_a) = n \log \left| \sum_{i=1}^{n} (\mathbf{y}_i - \hat{\mathbf{c}}(t_i) - \hat{\mathbf{g}}(t_i)) (\mathbf{y}_i - \hat{\mathbf{c}}(t_i) - \hat{\mathbf{g}}(t_i))^T \right| + 2dfc + 2dfg$$

where $\hat{\mathbf{c}}(\cdot)$ is the penalized fit with penalty parameters λ_a and $\hat{\mathbf{g}}(\cdot)$ is the fit based on $\tilde{\mathbf{y}}_t$ using the Mixed Model formulation from above. The degrees of freedom of the fits are calculated from

$$\begin{aligned} dfc &= \operatorname{tr}\left(\mathbf{Z}\left(\mathbf{Z}^T\mathbf{Z} + \mathbf{D}_a(\lambda_a)\right)^{-1}\mathbf{Z}^T\right) \text{ and} \\ dfg &= \operatorname{tr}\left(\mathbf{I}\left(\hat{\mathbf{b}}, \hat{\lambda}_b\right)^{-1}\mathbf{I}\left(\hat{\mathbf{b}}, \lambda_b = \mathbf{0}\right)\right). \end{aligned}$$

3 Application

3.1 Simulation

We demonstrate the performance of our routine with a small simulation study. We simulate data on a circle, i.e. $\tilde{y}_{t1} = \sin(t2\pi) + \varepsilon_{t1}$ and $\tilde{y}_{t2} = \cos(t2\pi) + \varepsilon_{t2}$, where $\varepsilon_{t1}, \varepsilon_{t2}$ are independent $N(0, 0.25^2)$ residuals and t ranges from 0 to 1 in $n = 200$ equidistant steps. The short term fluctuation is overlayed with a long term trend so that $y_{t1} = \tilde{y}_{t1} + 0.5\cos(10\pi t)$ and $y_{t2} = \tilde{y}_{t2} + 0.5\sin(10\pi t)$. We have used in all numerical estimations the same B-Spline Basis of order 3, i.e. $\mathbf{B}_\rho(t) = \mathbf{B}_\phi(t) = \mathbf{Z}_1(t) = \mathbf{Z}_2(t)$, and the same penalty matrices of order 2, i.e. $\mathbf{D}_\rho = \mathbf{D}_\phi = \mathbf{D}_a$. In Figure 2 we show simulated data and the corresponding estimates. The Figure, as well as all subsequent Figures, are organized as follows. The first two plots show the two time series y_{t1} and y_{t2}. The resulting long phase estimate is superimposed as solid line. The final estimate $\hat{\mathbf{c}}(t) + \hat{\mathbf{g}}(t)$ is shown with confidence bands for both series. The bottom row shows the observations (y_{t1}, y_{t2}) with the long term trend $\hat{\mathbf{c}}(t)$ (bottom left plot) and the residuals $(\tilde{y}_{t1}, \tilde{y}_{t2}) = (y_{t1} - \hat{\mathbf{c}}_1(t), y_{t2} - \hat{\mathbf{c}}_2(t))$ with the fitted shorter phase structure $\hat{\mathbf{g}}(t)$. Finally, the two right hand side plots show the fitted radius $\hat{\rho}(t)$ (upper right hand side plot) and the fitted angle $\hat{\phi}(t)$ (bottom right hand side plot). The separation of long term and short term fluctuation seems adequate for the data. The smoothing parameters for the long term trend are thereby selected following the Akaike criterion proposed above. The corresponding

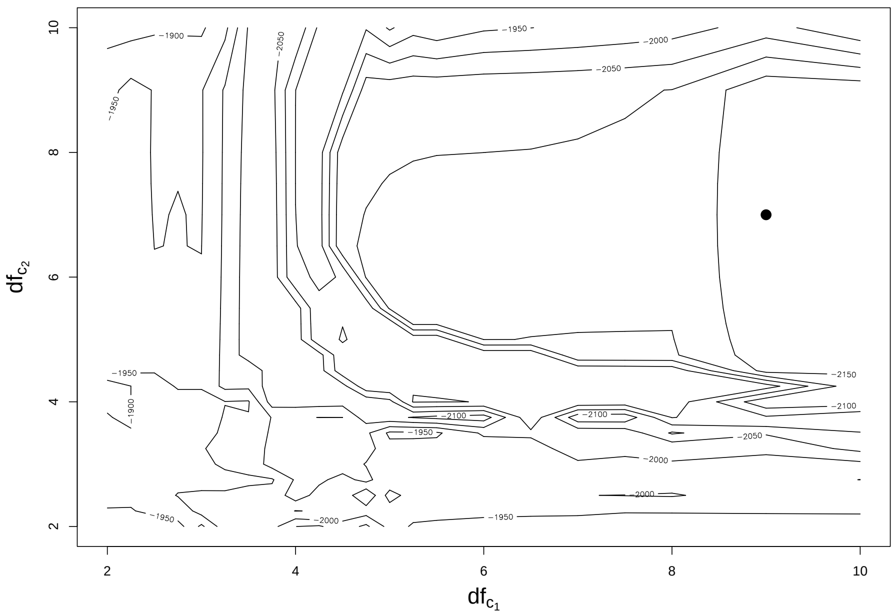

Figure 1: Shape of the AIC function for resulting degrees of freedom for the smoothing parameter of the long phase cycle

shape of AIC(λ_a) is shown exemplary in Figure 1. Clearly the existence of two phases is visible in the AIC function, where the value of AIC is given for the resulting degrees of freedom for $c_1(t)$ and $c_2(t)$. We can also see that the angle is estimated nearly linearly, indicating a constant rotation speed for $\mathbf{g}(t)$. In the same way, the radius is about constant over time. Apparently, the true function is reproduced for this simulation.

3.2 Business and long phase cycles in inflation and income distribution

Econometric studies often focus on the methodological level as well as in empirical research on the problem of how to separate the business cycle from the trend in important macroeconomic time series. Yet, economic growth theory in its advanced form provides us with insights on which economic ratios may exhibit a secular trend (like capital intensity when not measured in efficiency units) and which ones will not (like the output-capital ratio or the rate of employment as two measures of macroeconomic factor utilization). In contrast to a variety of econometric studies, macrodynamic growth theory therefore generally uses appropriate ratios or growth rates in its analytical investigations. In particular ratios are used that allow for the determination of steady state positions and which therefore should not exhibit a trend in the very long run.

In applying the methodology developed in this paper we will in fact concentrate on secularly trendless magnitudes, namely the employment rate on the external labor market, the wage share in national income and the inflation rate (here of producers' prices). There are a variety of smaller as well as larger macrodynamic models in the tradition of Friedman (1968) and Goodwin (1967) which show the existence of persistent cycles in the interaction between the employment rate and the wage share on the one hand and the employment rate and the inflation rate on the other hand which tend to long phased when simple constant parameter estimates are used for their numerical investigation (see also Atkinson 1969). In these models the ordinary business cycle fluctuations must therefore be explained by something else, namely by systematic variations in the parameters of the model which then add cycles of period lengths of about eight years to the fifty years cycles these models are generating when used with average or constant parameter values. Based on earlier work (see Flaschel, Kauermann, and Teuber 2006) we now investigate the working hypothesis that there are long phase cycles interacting with business cycles in the data as far as employment, income distribution and inflation are concerned. The method developed in this paper now in fact allows us to ceck this hypothesis in a way much more refined then just by using the Hodrick-Prescott filters with an arbitrarily given λ parameter. Moreover, we pursue the hypothesis in spirit of the two-dimensional phase plots of the employment–inflation cycle and the employment–income distribution cycle of the literature on the Friedman inflation cycle and the Goodwin growth cycle. Applying the technique developed leads to the estimates show in Figure 3 and Figure 4. We first focus on inflation dynamics that is Figure 3. We see that the unemployment rate is leading compared to the inflation rate in the long phase cycle (the solid lines in the two time series plots top-left). In the bottom figure showing angle estimate $\hat{\phi}(t)$ we see moreover that there are approximately 6 business cycles surrounding these long phase cycles, as $\hat{\phi}(t)$ crosses about 6 times the 2π full circle, marked as horizontal dashed lines. This finding is in line with Chiarella, Flaschel, and Franke (2005) and other work. The fitted angle also shows that the anti-clockwise rotation of the long phase cycle is by and large also characterized by the business cycles surrounding it, though there are exceptions to this rule (periods at the beginning and the end of the considered time span), see also the figure top-right. We note that we follow the tradition here which uses the unemployment rate in place of the employment rate on the horizontal axis (the latter would give rise to an anti-clockwise orientation of the business and the long phase cycles shown in these figures). The long phase cycle (bottom left plot of Figure 3) indicates that indeed 50 years of data are needed in order to get the indication of the existence of such a cycle. We observe that long periods where unemployment and inflation are both rising (i.e., where stagflation occurs) and also periods where the opposite takes place and therefore falling unemployment rates do not lead to rising inflation rates immediately. We stress again that our extraction of the business cycle component as shown in Figure 3 through a phase as well as a radius plot is an integral part of our treatment of the long phase evolution of the economy. With respect to the other long phase cycle model, the Goodwin (1967) growth cycle model, we have now look at Figure 4. As far as the evolution of the wage share (top-left plot) is concerned we have now more volatility as was the case with the inflation rate. This may be due to the involvement of labor productivity as constituent part of the definition of the wage share. Nevertheless one can see a single long phase cycle in the solid line shown in the time series presentation of the wage share. Again, the employment rate is leading with respect to this long phase cycle in the wage share. We know from Goodwin (1967) and the numerous articles that followed his approach that the interaction of the employment rate with the wage share is generating a clockwise motion. In Figure 4 we can in this regard confirm that the cycles of business cycle frequency are moving in a clockwise fashion as it is suggested by the again basically downward sloping angle line bottom-right. To the right of this figure we again see (if minor cycles are neglected) now by and large 7 business cycles overlaid over the long phase cycles as they are also shown in the figure bottom-right. Looking at the long-phase cycle (bottom-middle plot) we see indeed a cycle that is nearly closed (and thus approximately of fifty years length) and that is moving clockwise as suggested by the simple Goodwin (1967) growth cycle model (see Solow 1990 for early comments on an empirical phase plot of this cycle) and its many extensions. We conclude that the method developed in this paper provides a helpful approach to the separation of long-phased cycles that describe the evolution from high to low inflation regimes and from high to low wage share regimes from cycles of business cycle frequency. This method therefore allows in a distinct way the discussion of long waves in inflation and income distribution in modern market economies after World War II.

4 Conclusions

In the paper we attempted to fit and visualize long phase and business cycles. The intention was to treat cycles in a bivariate form and separation between long and short phase was possible by the use of nonparametric smoothing techniques. Particularly, the use of an Akaike criterion helped to decompose the two time trends. The underlying technique was built on P-spline smoothing which proves as flexible and general estimation tool. The plots accompanying the fit allow for empirical insight in business cycle theory. The joint estimation of the long phase cycle and business cycles confirm the classical economic models. The results showed that the Friedman and Goodwin cycles take about 50 years and are superposed by the business cycles with an approximately length of 8 years.

References

Akritas, M. and D. Politis (2003). *Recent Advances and Trends in Nonparametric Statistics*. Holland: North.

Atkinson, A. (1969). The timescale of economic models: How long is the long run? *Review of Economic Studies 36*, 137–152.

Backus, D., P. Kehoe, and F. Kydland (1992). International real business cycles. *Journal of Political Economy 100*(4), 745–775.

Baxter, M. and R. King (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *The Review of Economics and Statistics 81*(4), 575–593.

Breslow, N. E. and D. G. Clayton (1993). Approximate inference in generalized linear mixed model. *Journal of the American Statistical Association. 88*, 9–25.

Brockwell, P. J. and R. A. Davis (1987). *Time Series: Theory and Methods*. Springer-Verlag, Berlin, New York.

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

Chiarella, C., P. Flaschel, and R. Franke (2005). *Foundations for a Disequilibrium Theory of the Business Cycle. Qualitative Analysis and Quantitative Assement.* Cambridge: Cambridge University Press.

de Boor, C. (1978). *A Practical Guide to Splines*. Berlin: Springer.

Eilers, P. and B. Marx (1996). Flexible smoothing with b-splines and penalties. *Statistical Science 11*(2), 89–121.

Einbeck, J., G. Tutz, and L. Evers (2005). Local principle curves. *Statistics and Computed*, (conditionally accepted).

Eubank, R. L. (1989). *Spline smoothing and nonparametric regression.* Dekker: New York.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modelling and its Applications.* London: Chapman & Hall.

Fan, J. and Q. Yao (2003). *Nonlinear Time Series: nonparametric and parametric methods.* New York: Springer.

Fisher, N. (1995). *Statistical analysis of circular data* (3rd ed.). Cambridge: Cambridge University Press.

Flaschel, P., G. Kauermann, and T. Teuber (2006). Long cycles in employment, inflation and real unit wage costs. *American Journal of Applied Sciences*. forthcoming.

Friedman, M. (1968). The role of monetary policy. *American Economic Review 58*, 1–17.

Goodwin, R. (1967). A growth cycle. In C. Feinstein (Ed.), *Socialsim, Capitalism and Economic Growth*, pp. 54–58. Cambridge: Cambridge University Press.

Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*(2), 357–384.

Hamilton, J. (2005). What's real about the business cycle? *Federal Reserve Bank of St. Louis, Review 87*(4), 435–452.

Härdle, W. and J. S. Marron (1991). Bootstrap simultaneous error bars for nonparametric regression. *The Annals of Statistics 19*(2), 778–796.

Hastie, T. and W. Stützle (1989). Principle curves. *Journal of the American Statistical Association 84*, 502–516.

Hastie, T. and R. Tibshirani (1990). *Generalized Additive Models*. London: Chapman and Hall.

Hodrick, R. and E. Prescott (1981). Post-war u.s. business cycles: An empirical investigation. Discussion Paper 451, Northwestern University.

Hodrick, R. and E. Prescott (1997). Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking 29*, 1–16.

Härdle, W., H. Lütkepohl, and R. Chen (1997). A review of nonparametric time series analysis. *International Statistical Review 65*, 49–72.

Kauermann, G. (2004). Penalised spline fitting in multivariable survival models with varying coefficients. *Computational Statistics and Data Analysis*, (to appear).

Koopmans, T. C. (1947, August). Measurement without theory. *Review of Economic Statistics 29*, 161–172.

Krivobokova, T. and G. Kauermann (2004). A short note on penalized spline smoothing with correlated errors. Technical report.

Kydland, F. and E. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica 50*(6), 1345–1370.

Kydland, F. E. and E. C. Prescott (1990). Business cycles: Real facts and a monetary myth. *Federal Reserve Bank of Minneapolis Quarterly Review Spring*, 3–18.

Lindstrom, M. and D. Bates (1990). Nonlinear mixed-effects models for repeated measures data. *Biometrics 46*, 673–687.

Long, J. and C. Plosser (1983). Real business cycles. *Journal of Political Economy 91*(1), 39–69.

Lotka, A. (1925). *Elements of physical biology.* Baltimore: Williams & Wilkins Co.

Lucas, R. E. J. (1977). Understanding business cycles. In K. Brunner and A. H. Meltzer (Eds.), *Stabilization of the domestic and international economy*, Volume 5 of *Carnegie-Rochester Conference Series on Public Policy*, pp. 7–29. Amsterdam: North-Holland.

Mao, W. and L. Zhao (2003). Free-knot polynomial splines with confidence intervals. *Journal of the Royal Statistical Society, Series B 65*, 901–919.

Opsomer, J., Y. Wang, and Y. Yang (2001). Nonparametric regression with correlated errors. *Statistical Science 16*, 134–153.

O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems (c/r: P519-527). *Statistical Science 1*, 502–518.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics 11*, 735–757.

Ruppert, D. (2004). *Statistics and Finance - An Introduction*. New York: Springer.

Ruppert, D., M. Wand, and R. Carroll (2003). *Semiparametric Regression*. Cambridge University Press.

Sinclair, T. (2006). Asymmetry in the business cycle: New support for friedman's plucking model. Working paper, George Washington University.

Solow, R. (1990). Goodwin's growth cycle: Reminiscence and rumination. In K. Velupillai (Ed.), *Nonlinear and multisectoral macrodynamics. Essays in Honour of Richard Goodwin*, pp. 31–41. London: Macmillan.

Stock, J. H. (1987). Measuring business cycle time. *Journal of Political Economy 95*(6), 1240–1261.

Stock, J. H. and M. W. Watson (1999). Business cycle fluctuations in us macroeconomic time series. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1A, pp. 3–64. Elsevier.

Volterra, V. (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Mem. R. Accad. Naz. dei Lincei*.

Wahba, G. (1990). *Spline Models for Observational Data*. Siam.

Wand, M. (2003). Smoothing and mixed models. *Computational Statistics 18*, 223–249.

A Technical Details

A.1 Estimation and Properties of Estimates

We reformulate the model 2.4 by defining $\mathbf{B}(t) = \text{diag}(\mathbf{B}_\rho(t), \mathbf{B}_\phi(t))$ with $\text{diag}(\cdot)$ as block diagonal matrix. This yields the linear predictor as $\boldsymbol{\eta}(t) = \mathbf{B}(t)\mathbf{b}$. We denote the derivative of $\mathbf{g}(\cdot)$ with respect to $\boldsymbol{\eta}$ by

$$\nabla\mathbf{g}(t) = \rho(t)\begin{pmatrix} \cos\phi(t) & \sin\phi(t) \\ -\sin\phi(t) & \cos\phi(t) \end{pmatrix}. \tag{A.1}$$

This allows to write the first derivative of (2.8) as

$$\frac{\partial l_P\left(\mathbf{b},\lambda_b\right)}{\partial\mathbf{b}} = \sum_{i=1}^{n}\mathbf{B}^T(t_i)\nabla\mathbf{g}(t_i)\boldsymbol{\Sigma}_\epsilon^{-1}\left(\mathbf{y}_{t_i} - \mathbf{g}(t_i)\right) - \mathbf{D}(\lambda_b)\mathbf{b} = \mathbf{0}, \tag{A.2}$$

where $\mathbf{D}(\lambda_b)$ is a block diagonal of the form $\text{diag}\left(\lambda_\rho\mathbf{D}_\rho, \lambda_\phi\mathbf{D}_\phi\right)$. Solving $\partial l_P(\cdot)/\partial\mathbf{b} = \mathbf{0}$ provides the penalized estimate which can be calculated in the ordinary way using a Newton Raphson procedure (or Fisher scoring). Accordingly, the penalized Fisher matrix results to

$$\begin{aligned}\mathbf{I}(\mathbf{b},\lambda_b) &= -E\left[\frac{\partial^2 l_P(\mathbf{b},\lambda_b)}{\partial\mathbf{b}\partial\mathbf{b}^T}\right] \\ &= \sum_{i=1}^{n}\mathbf{B}^T(t_i)\nabla\mathbf{g}(t_i)\boldsymbol{\Sigma}_\epsilon^{-1}\nabla\mathbf{g}(t_i)^T\mathbf{B}(t_i) + \mathbf{D}(\lambda_b) \qquad \text{(A.3)}\end{aligned}$$

Moreover, we can use conventional likelihood theory to derive asymptotic properties for $\hat{\mathbf{b}}$, keeping the penalty parameter fixed. Under the technical assumption that design points t_i at which observations are collected become dense the variance of the estimate is asymptotically given by

$$\text{Var}\begin{pmatrix}\hat{\hat{\rho}}(t_0) \\ \hat{\hat{\phi}}(t_0)\end{pmatrix} = \mathbf{B}(t_0)\mathbf{I}\left(\hat{\mathbf{b}},\lambda_b\right)^{-1}\mathbf{I}\left(\hat{\mathbf{b}};\lambda_b = \mathbf{0}\right)\mathbf{I}\left(\hat{\mathbf{b}},\lambda_b\right)^{-1}\mathbf{B}^T(t_0). \tag{A.4}$$

Based on this result, the variance for radius and angle estimates is then obtained by the delta rule

$$\mathrm{Var}\left(\hat{\rho}(t), \hat{\phi}(t)\right) \approx \hat{\mathbf{C}}(t)\mathrm{Var}\left(\hat{\tilde{\rho}}(t), \hat{\tilde{\phi}}(t)\right)\hat{\mathbf{C}}(t)^T$$

with $\mathbf{C}(t) = \mathrm{diag}\left(\exp\left(\hat{\tilde{\rho}}(t)\right), 1\right)$, where, as motivated above, we ignored the mod$(\cdot)$ function. Accordingly, the variance for the cycle estimate $\mathbf{g}(t)$ results as $\mathrm{Var}(\hat{\mathbf{g}}(t)) \approx \hat{\mathbf{G}}(t)\mathrm{Var}(\hat{\tilde{\rho}}(t), \hat{\tilde{\phi}}(t))\hat{\mathbf{G}}(t)^T$ with

$$\hat{\mathbf{G}}(t) = \hat{\rho}(t)\begin{pmatrix}\cos\hat{\phi}(t) & -\sin\hat{\phi}(t)\\ \sin\hat{\phi}(t) & \cos\hat{\phi}(t)\end{pmatrix}.$$

The results follow directly in the line of Ruppert, Wand, and Carroll (2003).

A.2 Generalized Linear Mixed Models

It has shown to be advantageous, both in terms of numerics and theory, to link spline smoothing with linear mixed models. For P-spline smoothing this connection has been demonstrated in Wand (2003). We extend this idea here by formulating the penalization as an *a priori* distribution on the spline coefficients. This is available with the following reformulation:

$$\mathbf{y}_t|\mathbf{b} \sim N\left(\mathbf{g}(t), \boldsymbol{\Sigma}_\epsilon\right) \qquad \mathbf{L}\mathbf{b} \sim N\left(\mathbf{0}, \boldsymbol{\Lambda}_b\right) \tag{A.5}$$

with $\mathbf{g}(t)$ as defined in (2.3) and $\mathbf{L} = \mathrm{diag}\left(\mathbf{L}_\rho, \mathbf{L}_\lambda\right)$ and $\boldsymbol{\Lambda}_b = \boldsymbol{\Lambda}_b(\lambda_b)$ as block diagonal having $I_{k_\rho}\lambda_\rho^{-1}$ and $I_{k_\phi}\lambda_\phi^{-1}$ on the diagonal where k_ρ and k_ϕ are the dimensions of spline bases $\mathbf{B}_\rho$ and $\mathbf{B}_\phi$, respectively. Now, penalty parameter $\lambda = (\lambda_\rho, \lambda_\phi)$ expresses the *a priori* precision, that is $1/\lambda$ gives the a priori variance for spline coefficients treated as random coefficients. Integrating out $\mathbf{Lb}$ we obtain the marginal log likelihood based on the mixed model (A.5)

$$l_{mm}\left(\boldsymbol{\Sigma}_{\epsilon}, \lambda_b\right) = \log \int \frac{1}{|\boldsymbol{\Lambda}_b|^{1/2}} \exp\{l_p\left(\mathbf{b}, \boldsymbol{\Sigma}_{\epsilon}; \lambda_b\right)\} d\mathbf{Lb} \tag{A.6}$$

The objective is now to maximize (A.6) with respect to λ_b and $\boldsymbol{\Sigma}_{\varepsilon}$ and predict the spline coefficients $\mathbf{b}$ to achieve a smooth fit. Note that maximization with respect to λ_b provides an estimate for the amount of penalization required. Apparently, due to the non linear link used for the mean structure, (A.6) does not yield an analytic solution. Instead, a Laplace approximation can be used in the line of Breslow and Clayton (1993) or Lindstrom and Bates (1990). It is shown in the Appendix that the penalized fit from above is equivalent to a posterior mode estimate in the mixed model. Moreover, smoothing parameter $\lambda = (\lambda_\rho, \lambda_\phi)$ can be chosen to maximize the (Laplace approximated) marginal likelihood. In particular we get the estimate

$$\frac{1}{\hat{\lambda}_\rho} = \frac{\mathrm{tr}\left\{\left(\mathbf{I}(\boldsymbol{b}, \lambda)^{-1}\right)_{\rho\rho} \mathbf{D}_\rho\right\} + \hat{\mathbf{b}}_\rho^T \mathbf{D}_\rho \hat{\mathbf{b}}_\rho}{k_\rho} \tag{A.7}$$

Finally, based on the Laplace approximation likelihood an estimate for $\boldsymbol{\Sigma}_\epsilon$ is defined through

$$\hat{\boldsymbol{\Sigma}}_\epsilon = \frac{\sum_{i=1}^{n} \{\mathbf{y}_i - \hat{\mathbf{g}}(t_i)\}\{\mathbf{y}_i - \hat{\mathbf{g}}(t_i)\}^T}{n} + O(n^{-1}). \tag{A.8}$$

A.3 Numerical and Practical Adjustments

Confidence regions

For each timepoint we obtain estimates and confidence intervals for the fitted functions $g_1(t)$ and $g_2(t)$. We are however more interested in confidence regions for the fitted two dimensional curves $(g_1(t), g_2(t))$. These are achieved using the asymptotic arguments from above and constructing a confidence ellipse at timepoint t through

$$\mathrm{CI}(t) \quad = \quad \left\{\mathbf{z} : (\mathbf{z} - \hat{\mathbf{g}}(t))^T \mathrm{Var}(\hat{\mathbf{g}}(t))^{-1} (\mathbf{z} - \hat{\mathbf{g}}(t)) \leq \chi^2_{2,\,0.95}\right\}$$

with $\chi^2_{2,\,0.95}$ as 95% Quantile of a χ^2 distribution with 2 degrees of freedom. One should note that the confidence ellipse are constructed pointwise and a global confidence level for CR is therefore not easily available. This is, however, a standard problem in smoothing. Moreover, the confidence ellipse does not mirror the variability due to the estimation of the smoothing parameter. For simplicity, we do not investigate these two issues in more depth here (see also Mao and Zhao 2003, or Härdle and Marron 1991 for a more theoretical consideration of these points).

Correlated Errors

For time dependent data it is generally difficult to distinguish between trend and correlation. For P-spline smoothing it has been shown in Krivobokova and Kauermann (2004) that residual correlation in a normal smoothing model does only have a weak influence on the resulting fitted trend of Maximum Likelihood or Residual based Maximum Likelihood (REML) smoothing parameter selection. We conjecture that this result also holds for the non-normal model fitted here, but we have no formal proof at hand. Instead, we exemplify the point with some simulations below. In general, of course, a unique decomposition of trend and correlation is impossible. It should also be noted that in principle a two step fitting can be pursued. First, a mean structure can be fitted which is then used to estimate the temporal correlation from the residuals. This is again used to recalculate both, the fit as well as the smoothing parameter. We do not go further this road, but explore an Akaike criterion instead.

A.4 Laplace Approximation

We derive the equivalence between penalized spline smoothing with B-splines and mixed models for the simple smoothing model

$$\mathrm{E}\left(y|t\right) = g\left(\mathbf{B}(t)\mathbf{b}\right)$$

with $\mathbf{B}(t)$ as B-spline basis of dimension k and order p, say. Extensions to the bivariate fitting routine described in Section 2 are straight forward. Let $\mathbf{u} := \mathbf{L}\mathbf{b}$ with $\mathbf{L}$ as difference matrix of order p, say, so that $\mathbf{L} \in \mathbb{R}^{k\times(k-p)}$. We can complete $\mathbf{L}$ by adding linearly independent rows such that

$$\tilde{\mathbf{u}} = \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{L}_0 \\ \mathbf{L} \end{pmatrix} \mathbf{b} = \tilde{\mathbf{L}}\mathbf{b}$$

with $\tilde{\mathbf{L}}$ invertible. For fitting we now impose penalty $\lambda \mathbf{b}^T \mathbf{D} \mathbf{b}$ on $\mathbf{b}$, where $\mathbf{D} = \mathbf{L}^T \mathbf{L}$. This penalty will be comprehend as *a priori* normal distribution for $\mathbf{u} = \mathbf{L}\mathbf{b} \sim N(0, I_{(k-p)}/\lambda)$ with $I_{(k-p)}$. Integrating out $\mathbf{u}$ leaves us with the remaining (unpenalized) parameter $\mathbf{u}_0$. This leads to the marginal likelihood for $\mathbf{u}_0$ and λ given through

$$l_{mm}(\lambda, \mathbf{u}_0) = \log \int \exp\left(l_p(\mathbf{u}, \lambda)\right) d\mathbf{u}$$

with $l_p(\mathbf{u}, \lambda)$ as penalized log likelihood defined through

$$\log \left\{ f(y|\mathbf{u})\, \varphi\left(\mathbf{u}, I_{(k-p)}/\lambda\right) \right\}$$

where $f(y|u)$ is the density of y given the linear predictor $\mathbf{B}(t)\mathbf{b} = \mathbf{B}(t)\tilde{\mathbf{L}}^{-1}\tilde{\mathbf{u}}$ and $\varphi(\cdot)$ is the normal density. Laplace approximation with respect to $\mathbf{u}$ and maximization with respect to the remaining parameter $\mathbf{u}_0$ leads to the maximized marginal log likelihood

$$\begin{aligned} l_{mm}(\lambda, \hat{\mathbf{u}}_0) &\approx l_p\left(\hat{\mathbf{b}}, \lambda\right) - \frac{1}{2}\log\left|\frac{I_{(k-p)}}{\lambda}\right| \\ &\quad -\frac{1}{2}\log\left|\left(0, I_{(k-p)}\right)(\tilde{\mathbf{L}}^{-1})^T \frac{\partial^2 l_p\left(\hat{\mathbf{b}}, \lambda\right)}{\partial \mathbf{b}\partial \mathbf{b}^T}\tilde{\mathbf{L}}^{-1}\begin{pmatrix} 0 \\ I_{(k-p)} \end{pmatrix}\right| \\ &\approx l_p(\hat{\mathbf{b}}, \lambda) - \frac{1}{2}\log\left|\mathbf{F}\left(\hat{\mathbf{b}}, \lambda\right)\frac{I_{(k-p)}}{\lambda}\right| \end{aligned} \tag{A.9}$$

with

$$\mathbf{F}(\mathbf{b}, \lambda) = \left(0, I_{(k-p)}\right)(\tilde{\mathbf{L}}^{-1})^T \mathbf{I}(\mathbf{b}, \lambda)\, \tilde{\mathbf{L}}^{-1}\begin{pmatrix} 0 \\ I_{(k-p)} \end{pmatrix}.$$

The latter simplification results by replacing the second order derivative by the Fisher information (A.3). It is also not difficult to show that $\hat{\mathbf{b}}$ is the solution to the penalized score (see (A.2)) with $\hat{\mathbf{u}} = \left(\hat{\mathbf{u}}_0^T, \hat{\mathbf{u}}^T\right) = \tilde{\mathbf{L}}\hat{\mathbf{b}}$. Differentiating (A.9) with respect to λ yields

$$0 = -\hat{\mathbf{b}}^T \mathbf{L}^T \mathbf{L} \hat{\mathbf{b}} + \frac{(k-p)}{\lambda} - \operatorname{tr}\left(\mathbf{F}^{-1}\left(\hat{\mathbf{b}}, \lambda\right)\right).$$

To simplify numerics and since $\mathbf{F}^{-1}(\mathbf{b}, \lambda)$ is of order $O(n^{-1})$ we approximate the formula with

$$\operatorname{tr}\left(\mathbf{F}^{-1}(\mathbf{b}, \lambda)\right) \approx \operatorname{tr}\left(\left((\tilde{\mathbf{L}}^{-1})^T \mathbf{I}(\mathbf{b}, \lambda)\, \tilde{\mathbf{L}}^{-1}\right)^{-1}\right) = \operatorname{tr}\left(\mathbf{I}^{-1}(\mathbf{b}, \lambda)\, \mathbf{L}^T \mathbf{L}\right)$$

which suggests the estimating equation

$$\hat{\lambda} = \frac{\hat{\mathbf{b}}^T \mathbf{D} \hat{\mathbf{b}} + \operatorname{tr}\left(\mathbf{I}^{-1}\left(\hat{\mathbf{b}}, \lambda\right) \mathbf{D}\right)}{k-p} \tag{A.10}$$

with $\mathbf{D} = \mathbf{L}^T \mathbf{L}$. It is thereby worth pointing out that (A.10) does not provide an analytic solution since the right hand side also depends on λ_ρ through $\hat{\mathbf{b}}_\rho$. Equation (A.10) can however be used in a backfitting style. This means we estimate $\hat{\mathbf{b}}$ through (A.2) by keeping λ fixed. Next, we consider $\hat{\mathbf{b}}$ as fixed and update λ through (A.10). Iteration between these two steps mirrors the backfitting iterations. We also refer to Krivobokova and Kauermann (2004) for a justification of this algorithm as Newton procedure.

B B-Splines

A general construction principle for B-splines is found in de Boor (1978) and implemented in the Splus function bspline($\cdot$). We work with B-splines of order three which are for equidistant grid point defined through

$$B_l(t) = \begin{cases} \frac{(t-\tau_{l-4})^3}{6h^3} & \text{for } t \in [\tau_{l-4}, \tau_{l-3}] \\ 2/3 - \frac{(t-\tau_{l-2})^2}{h^2} - \frac{(t-\tau_{l-2})^3}{2h^3} & \text{for } t \in [\tau_{l-3}, \tau_{l-2}] \\ 2/3 - \frac{(t-\tau_{l-2})^2}{h^2} + \frac{(t-\tau_{l-2})^3}{2h^3} & \text{for } t \in [\tau_{l-2}, \tau_{l-1}] \\ \frac{-(t-\tau_l)^3}{6h^3} & \text{for } t \in [\tau_{l-1}, \tau_l] \\ 0 & \text{otherwise.} \end{cases}$$

where $\tau_{-3} = \tau_{-2} = \tau_{-1} = \tau_0$ and $h = \tau_j - \tau_{j-1}$ is the distance between two neighbor knots.

Accordingly we use a second order difference for penalization, that is

$$L = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & \cdots \\ 0 & 1 & -2 & 1 & 0 & \cdots \\ & & \vdots & & & \ddots \end{pmatrix}.$$

C Simulating and Estimating the model

C.1 Simulation studies

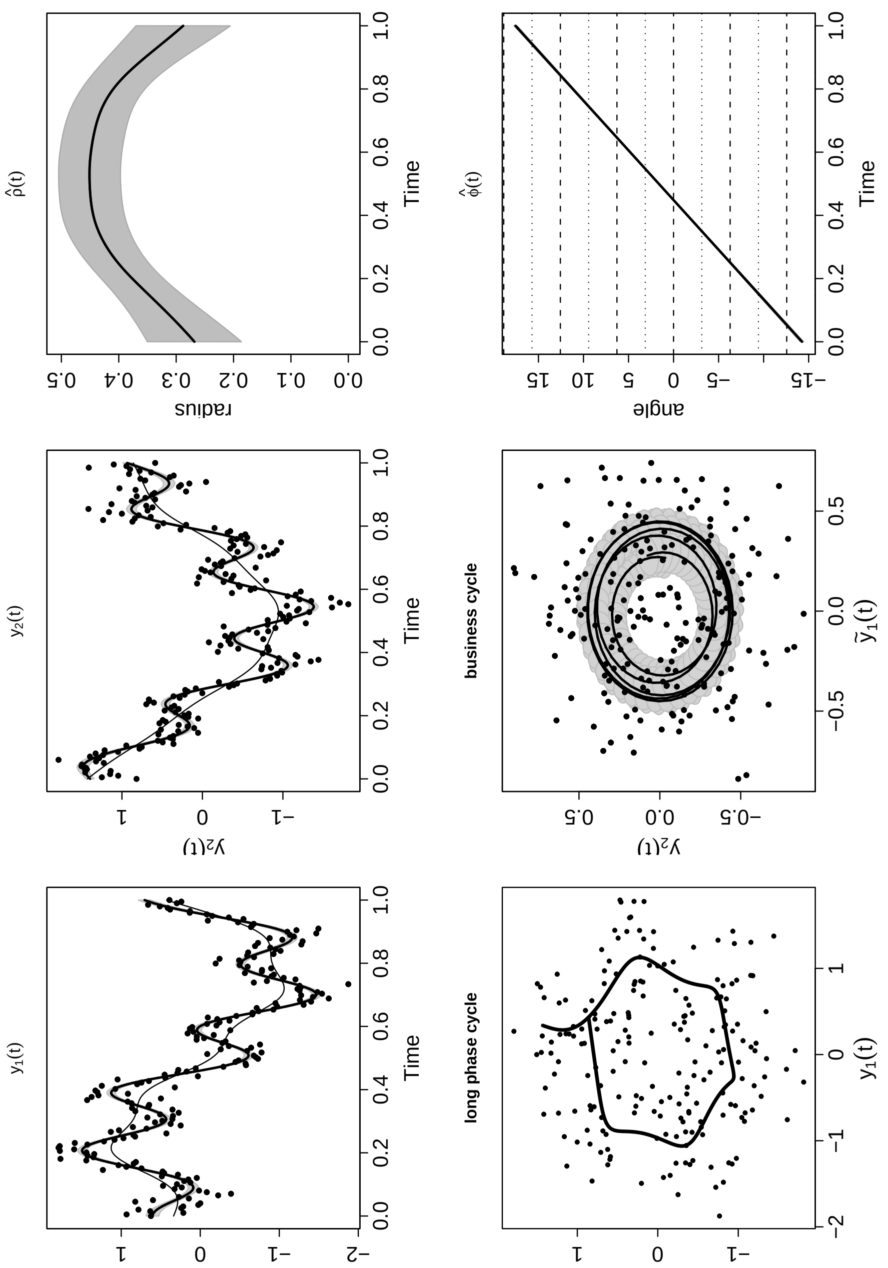

Figure 2: Upper Row: The left and middle plot show the artificial time series y_{t1} and y_{t2}, with long phase estimate (solid line), final estimation (bold solid line in the grey shaded area) with confidence region. The right plot shows the estimated radius and its confidence region (grey shaded area). Lower Row: The left plot shows the observations with the estimated long phase cycle (solid line). The middle plot shows the detrended time series with estimated business cycle (solid line) and its confidence region (grey shaded area). The right plot shows the estimated angle with its confidence region (grey shaded area).

C.2 Business cycles around long phase cycles: The US Economy

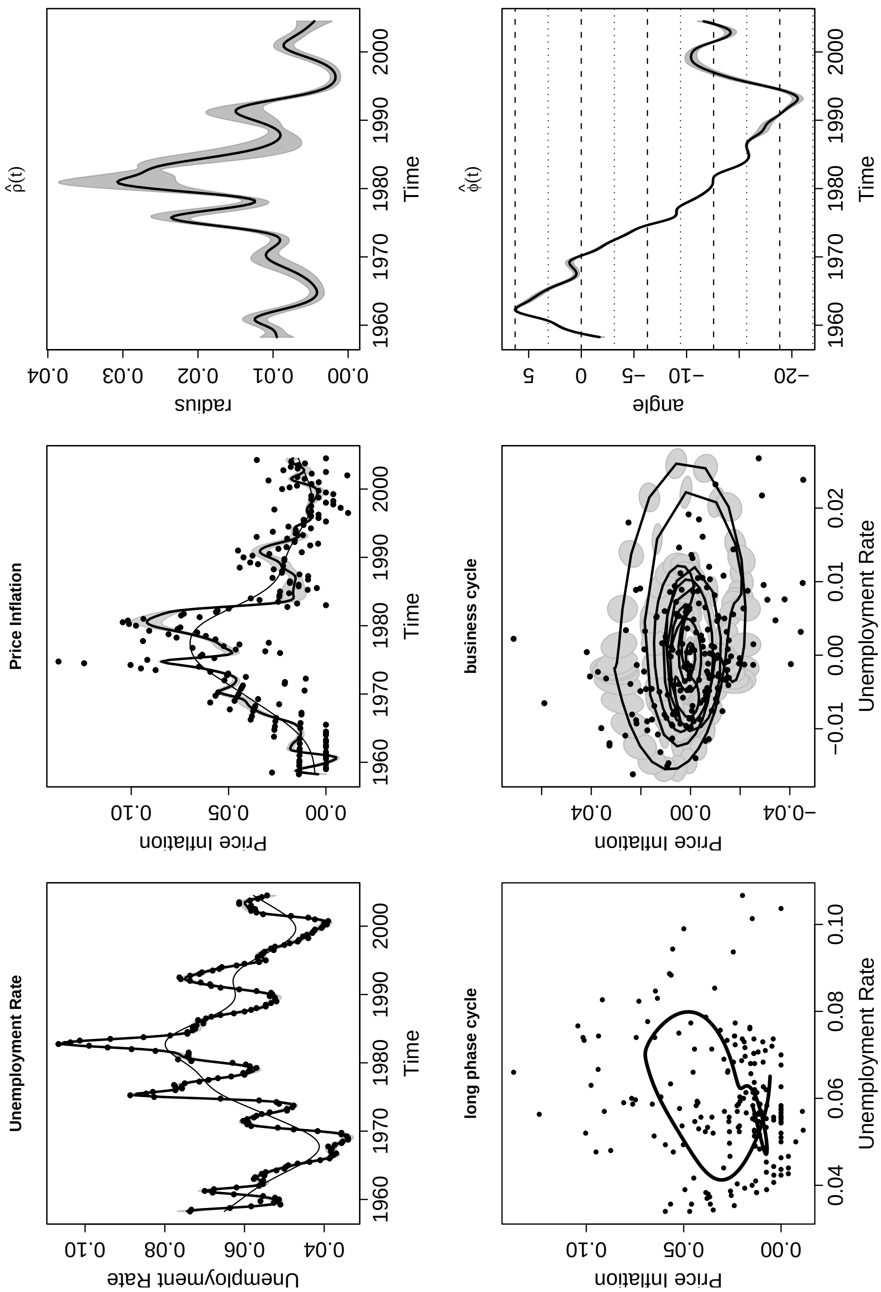

Figure 3: Upper Row: The left and middle plot show the time series Unemployment Rate (y_{t1}) and Price Inflation (y_{t2}), with long phase estimate (solid line), final estimation (bold solid line in the grey shaded area) with confidence region. The right plot shows the estimated radius and its confidence region (grey shaded area). Lower Row: The left plot shows the observations with the estimated long phase cycle (solid line). The middle plot shows the detrended time series with estimated business cycle (solid line) and its confidence region (grey shaded area). The right plot shows the estimated angle with its confidence region (grey shaded area).

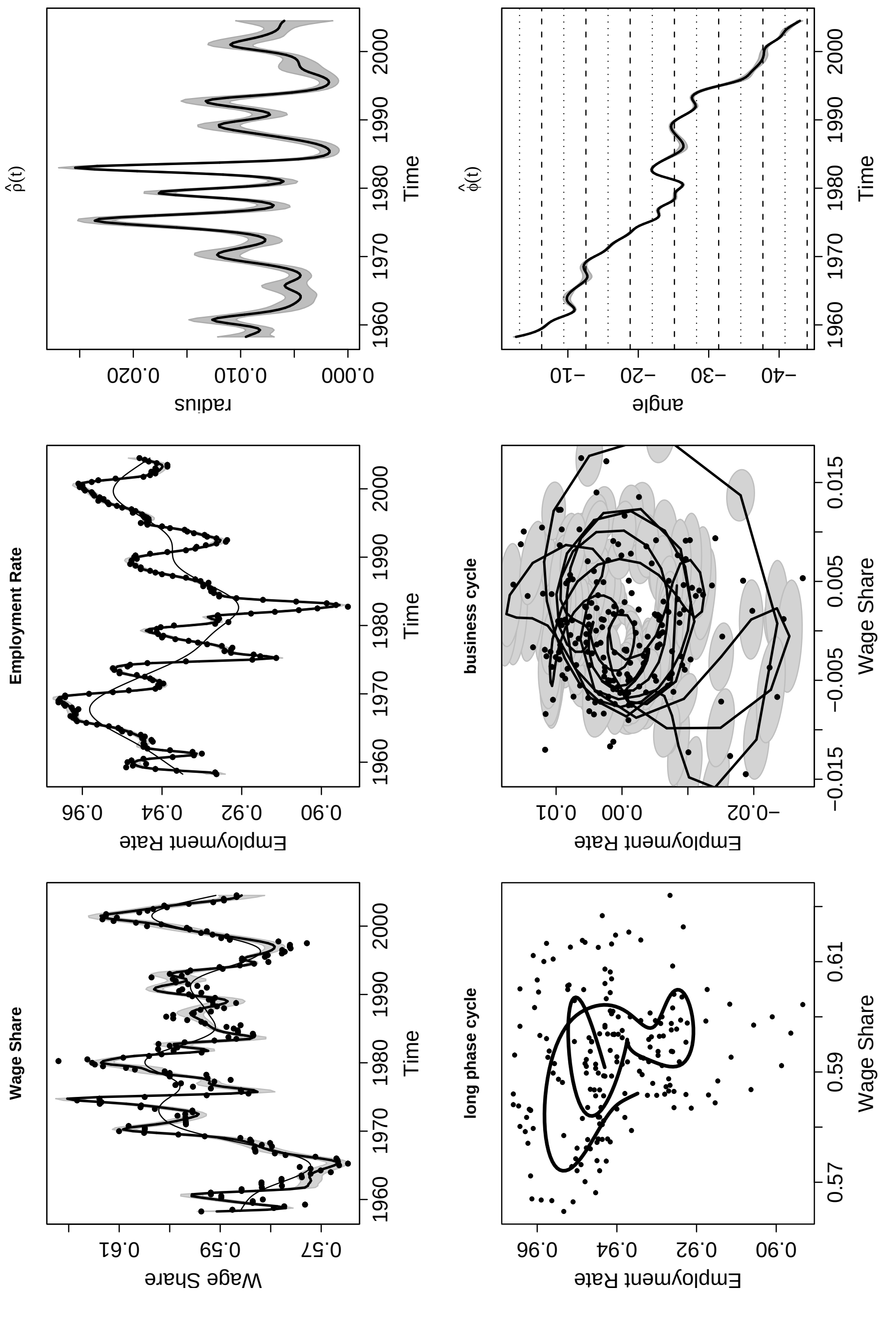

Figure 4: Upper Row: The left and middle plot show the time series Wage Share (y_{t1}) and Employment Rate (y_{t2}), with long phase estimate (solid line), final estimation (bold solid line in the grey shaded area) with confidence region. The right plot shows the estimated radius and its confidence region (grey shaded area). Lower Row: The left plot shows the observations with the estimated long phase cycle (solid line). The middle plot shows the detrended time series with estimated business cycle (solid line) and its confidence region (grey shaded area). The right plot shows the estimated angle with its confidence region (grey shaded area).