1 Introduction

The decomposition and filtering of time series is an important issue in economics and econometrics and related fields. Even though there are numerous competing methods on the market, in applications one often meets one of the few favorites. The first method to mention in this selection is the so called Hodrick and Prescott-filter (HP-filter hereafter). The idea is to decompose a time series y_t, say, into a smooth path g_t, also called non-stationary trend, and remaining deviations (residuals or business cycle components) ε_t which are assumed to be stationary around the trend. To achieve smoothness a penalty is imposed on g_t such that second order differences are penalized. The idea traces back to Leser (1961) and Whittaker (1923) and is simple in its numerical implementation, see Pedregal and Young (2001) for more general discussion on the the HP filter.

The application requires the specification of a penalty parameter λ, say, which steers the smoothness of the fitted path $\hat{g}_t$. Hodrick and Prescott (1997) refrain from suggesting any data driven choice for the amount of penalization, but develop a substance matter explanation. Understanding the smoothing parameter as a ratio of two variances describing short and long phase variability of y_t they argue to fix the penalty parameter λ at a given value ($\lambda = 1600$) so that quarterly changes are related to yearly variation of the smooth path g_t. This viewpoint changes if a different resolution of the data is considered like monthly or yearly observations. Ravn and Uhlig (2002) give a theoretical derivation how to adjust the penalty parameter λ in this case. In fact, following the economic interpretation of λ given in Hodrick and Prescott (1997) they suggest to take $\lambda = 1600/4^4$ for yearly and $\lambda = 1600 \cdot 3^4$ for monthly observations, respectively. Though this argumentation is sound and justifiable on economic grounds, it is weak following statistical thinking.

Schlicht (2005) suggests a data driven choice of the penalty parameter by understanding the penalty as prior distribution which leads to a so called Mixed Model, see also Dermoune, Djehiche, and Rahmania (2008). The penalty parameter is then the ratio of the residual variance and the variance of the *a priori* distribution which can be estimated in the Mixed Model framework. This result has also been shown in Harvey and Jaeger (1993) and is further explored in Dermoune, Djehiche, and Rahmania (2009). A Bayesian perspective in this direction has been proposed by Trimbur (2006) (see also (Harvey, Trimbur, and Van Dijk, 2007)). Both, the Mixed Model setting as well as the Bayesian approach assume that the remaining residuals are unstructured and without serial correlation which should be seen critically since it is not necessarily met in practice. Further critique concerning the HP filter has been formulated, e.g. by Cogley and Nason (1995) and Schenk-Hopp (2001).

The idea of imposing a prior distribution on the trend is in line with results derived for spline smoothing. In fact, the HP-filter can be comprehended as a spline smoother so that the penalty parameter becomes a smoothing parameter which could be estimated data driven following e.g. Wood (2000) or Hastie and Tibshirani (1990). In principle this formulation leads to a simple and feasible routine for selecting the smoothing parameter data driven. However, the method still relies on the crucial and questionable assumption that the remaining deviations ε_t are just white noise, and in particular not correlated. Incorporating correlated residuals into a cross validation criterion has been suggested in Kohn, Ansley, and Wong (1992) or Wang (1998). However, even if the correlation is incorporated, it can be demonstrated that even minor misspecifications of the correlation structure let available cross validation routines fail, as convincingly demonstrated in Opsomer, Wang, and Yang (2001), see also Proietti (2005) or Dagum and Giannerini (2006).

In compensation to this drawback Krivobokova and Kauermann (2007) show that the use of so called penalized splines in combination with Mixed Models provides a robustification against misspecified residual structure in the model. Penalized splines are thereby a relatively new smoothing technique which traces back to O'Sullivan (1986), see also Eilers and Marx (1996) or Ruppert, Wand, and Carroll (2003). The idea is to estimate the smooth component g_t by using a high but finite dimensional spline basis, and instead of simple parametric fitting one imposes a penalty on the spline coefficients, in close analogy to the HP filter. We give more details in the paper and show how the numerically simple method can be used for time series decomposition and might compete with the HP filter in a wide range of examples.

As mentioned before, the HP-filter and its extensions are just one of the available and commonly used candidates for time series decomposition. A further favorite method is the bandpass filter (BP-filter hereafter) and its different approximations and extensions, see Baxter and King (1999) , Christiano and Fitzgerald (2003) and Stock and Watson (1999). Here, the idea is to decompose the series y_t in its frequency domain. In fact, the intention is to decompose $y_t = g_t + \varepsilon_t$ where g_t has power in a prespecified frequency interval $(a,b) \in [0,2\pi]$. The BP-filter is constructed using a projection of y_t on the specified frequency range. It combines a high pass and low pass filter and has achieved quite some reputation in practice. In the BP-filter, apparently, the frequencies a and b play the role of smoothing parameters which influence the performance of the fit, where in practice economic considerations suggest its specification. Critique about the BP-filter has been formulated among other by Goldrian (2005) or Murray (2003). In this paper we will apply the BP filter in its original form to compare it with the proposed penalized spline estimate. Additionally we will reformulate the BP filter and write it as penalized spline fit. To do so we choose a rich dimensional basis covering the specified frequency domain and impose a suitable penalty on the spline coefficients. This connects the filter to the previously discussed extension of the HP-filter by just using a different basis.

An extensive overview on recent research on detrending and filtering and their use in economics is given in Canova (2007). Usually, economic time series data display trends and it is not immediately obvious what cyclical properties the data have. Since economists are interested in cyclical components and in the cross correlation of cyclical variables we need methods that separates trends from cyclical components. Detrending in economics is also needed in order to make economic time series stationary so that one can compute functions of second moments of the data. If one is, in business cycle analysis, only interested in other cyclical information, such as turning points of economic data, one does not necessarily need detrending. On the other hand filtering aims at much broader applications. In economics one often is only interested in filtering out low or high frequency components of the data or some harmonic oscillation generating some periodic movements through the use of sine and cosine functions. In Canova (2007) one can find a survey of such a broader class of filtering procedures and their respective advantages and short comings. In macroeconomics the major effort has been to decompose time series data in trend and cyclical components. For this purpose the most popular procedures have been the HP filter and the BP filter. We will thus concentrate our study of the comparison of the latter two filters with penalized splines as a procedure to obtain trend and cyclical components of macroeconomic time series data. Stock and Watson (1999) have decomposed seventy U.S. macroeconomic time series into trend and cyclical components. This work has had an extensive influence on the thinking of the cyclicity, the cross-correlation and the empirical regularities in the postwar U.S. time series data. We will pick up some of these examples and demonstrate the performance of the different filters.

The focus of our presentation is to decompose a time series into trend and deviations, with the latter is commonly called residuals or cyclical component or business cycle, respectively and denoted as ε_t, see e.g. Zarnowitz and Ozyildirim (2005). It is generally not plausible to assume that ε_t is just white noise, but maintaining or allowing for a serial correlation over ε_t is desirable. This is the starting point for our proposal and we demonstrate that the penalized spline approach can easily accommodate serial correlation. Contrary one must state that in the real business cycle literature the issue of serial correlation in the choice of the filter does not seem to have been sufficiently explored. This holds for the HP filter (see (Cooley, 1995), or more recently (Chari, Kehoe, and McGrattan, 2007) and (Christiano and Davis, 2006)) as well as for the BP-filter (see the extensive study of U.S. time series data (Stock and Watson, 1999)). We supplement this discussion by applying the penalized spline versions of the HP-and BP-filter to a number of time series listed in

Stock and Watson (1999) and discuss and explain resulting differences. The second contribution of this paper is to demonstrate the simple feasibility of the routines. In fact, we will take advantage of the open source software **R** and show that the filters are easily calculated which allows its investigation of their performance.

The paper is organized as follows. In Section 2 we describe penalized spline smoothing as general smoothing technique and relate this to HP- and BP-filtering, respectively. In Section 3 we apply the routines to a number of time series. We conclude and finalize the presentation in Section 4.

2 Penalized Spline Smoothing

2.1 Decomposing Time Series with Penalized Splines

We consider the time series $Y = (y_1, y_2, y_3, \ldots)^T$ and assume that y_t decomposes to

$$y_t = g_t + \varepsilon_t, \tag{1}$$

with ε_t as residual or unexplained short term variation and g_t as trend or long phase variation. The intention is to decompose y_t according to (1), that is to find a suitable filter to extract g_t from y_t. We propose to make use of penalized splines. To do so, let $B(t)$ denote a rich spline basis with support over the observed time points t. A simple possible choice is to use so called truncated polynomials in the form

$$B(t) = \left(1, t, \ldots, t^q, (t-\tau_1)_+^q, \ldots, (t-\tau_p)_+^q\right), \tag{2}$$

where q is the degree of the highest polynomial, $(t)_+ = t$ for $t > 0$ and $(t)_+ = 0$ otherwise. The knots $\tau_1, \ldots, \tau_p$ are equidistantly chosen covering the range of time points t. Practical choices for q are $q = 1$ or $q = 2$, respectively, and the knots τ_k may be places every 5th to 10th observation. Even though (2) is a convenient choice in practice, the approach of penalized spline smoothing is not restricted to any specific basis and other bases can be used including for instance a B-spline basis or radial basis functions (for more details see (Ruppert et al., 2003)). For now, however, (2) may be seen as one possibility and due to its simplicity it is further used to demonstrate the idea of penalized splines. We decompose basis $B(t) = \{X(t), Z(t)\}$ with $X(t)$ as low dimensional and $Z(t)$ as high dimensional part. For instance with $B(t)$ as given in (2) we set $X(t) = (1, t, \ldots, t^q)$ and $Z(t) = ((t-\tau_1)_+^q, \ldots, (t-\tau_p)_+^q)$. We now reformulate (1) to

$$y_t = B(t)\theta + \varepsilon_t = X(t)\beta + Z(t)u + \varepsilon_t, \tag{3}$$

with $\theta = (\beta^T, u^T)$ as coefficient vector. The residual vector $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots)^T$ is assumed to be normally distributed with some stationary correlation matrix R_ε, i.e.

$\varepsilon \sim N(0, \sigma_\varepsilon^2 R_\varepsilon)$. For fitting we impose a penalty on u leading to the penalized least square

$$l(\beta, u; h) = \{Y - B(t)\theta\}^T R_\varepsilon^{-1} \{Y - B(t)\theta\} + \frac{1}{2}\lambda u^T D u, \tag{4}$$

where D is a penalty matrix. For truncated polynomials the penalty matrix D is chosen as identity matrix I_p (see (Ruppert et al., 2003)). Finally, coefficient λ in (4) is the penalty parameter steering the amount of penalization. Setting $\lambda \to \infty$ gives a simple polynomial fit based on matrix $X(t)$ only while $\lambda \to 0$ yields an unpenalized fit based on the full basis matrix $B(t)$.

The important feature of penalized spline smoothing is its link to Mixed Models. Comprehending the penalty in (4) as a priori normal distribution and postulating normality for the remaining component ε_t leads to the linear Mixed Model of the form

$$Y|u \sim N\left(X\beta + Zu, \sigma_\varepsilon^2 R_\varepsilon\right), \quad u \sim N\left(0, \sigma_u^2 D^-\right), \tag{5}$$

with X and Z as design matrices built from rows $X(t)$ and $Z(t)$ with $t = 1, 2, 3, \ldots$, D^- as (generalized) inverse of D and smoothing coefficient $\lambda = \sigma_\varepsilon^2 / \sigma_u^2$. In (5), β as well as λ play the role of parameters which can be estimated with appropriate software for Mixed Models. We give details in the Appendix. This means the penalty parameter and the remaining parameters are available with well developed Maximum Likelihood theory for Mixed Models (see (Searle, Casella, and McCulloch, 1992)). In particular, Schall's algorithm 1991 can be used to estimate λ data driven in an iterative form. Moreover, the estimate for g_t results through $X(t)\hat{\beta} + Z(t)\hat{u}$ with $\hat{u}$ as posterior Bayes estimate or Best Linear Unbiased Predictor (BLUP). A useful feature of model (5) applied for smoothing is derived in Krivobokova and Kauermann (2007) where it is shown that Maximum Likelihood estimates are robust with respect to misspecification of R_ε. That is to say in case of serial correlation among ε_t model (5) shall provide a reasonable fit even if R_ε is not equal to the true (unknown) correlation structure of the residuals, as long as the misspecification is not too strong. This is an important advantage of penalized spline smoothing for detrending time series. What we will do in practice is to assume a simple autocorrelated process for ε_t, e.g. an AR(q) process, which yields a good estimate for g_t even if the true process does not follow exactly the specified serial correlation process (which it will never do). It is important to reflect that such a property does not hold for other smoothing and filtering techniques (see e.g. (Opsomer et al., 2001)) and therewith gives penalized spline smoothing a clear advantage which we make use of subsequently. Moreover, the Mixed Model representation of penalized splines allows estimation of the correlation matrix R_ε from the corresponding (log-) likelihood along with the other model parameters, as long as the general structure of R_ε

is specified, for instance as an AR(q) process. This can be done with any Mixed Model software available, as demonstrated in the Appendix.

2.2 Hodrick-Prescott Filter

The HP filter relies on model (1) where g_t is fitted by minimizing the penalized least square

$$\sum_t (y_t - g_t)^2 + \lambda \sum_{t \geq 3} \left\{ (g_t - g_{t-1}) - (g_{t-1} - g_{t-2}) \right\}^2, \tag{6}$$

with λ as a tuning parameter. Note that the fitting criterion (6) resembles the penalized least square (4) with the crucial simplification, that the residuals ε_t are simply white noise. It should be clear that this assumption is not necessarily realistic. Apparently, the HP filter is a penalized smoother with knots at each observed time point and penalty in form of the squared second order difference matrix. However, in contrast to a standard penalized smoothers, Hodrick and Prescott (1997) suggest to set $\lambda = 1600$ for quarterly data. Adjustment of the smoothing parameter to other data frequencies has been considered by Ravn and Uhlig (2002), who suggested $\lambda_s = s^4\, 1600$, with $s = 1/4$ for annual data and $s = 3$ for monthly. Using these λ values is economic theory based and hence ignores the information available from the data. In some instances this is critical and results in completely inappropriate decompositions, as discussed e.g. by Canova (1998) or Schenk-Hopp (2001). Schlicht (2005) suggests to use the link to Mixed Models to estimate the smoothing parameter data driven. However, this would fail, if the assumption of independent residuals is violated, as already noticed in the previous section. That is to say if ε_t does not mirror simple white noise but contains serial correlation, the selected smoothing parameter will be unsatisfactory.

2.3 The Bandpass Filter

We understand y_t now as resulting from the frequency process

$$y_t = \int_0^1 \left\{ c(\omega)\cos(2\pi\omega t) + d(\omega)\sin(2\pi\omega t) \right\} d\omega,$$

with $c(\omega)$ and $d(\omega)$ as weight functions. The idea of the BP filter is to decompose y_t to

$$y_t = g_t + \varepsilon_t, \tag{7}$$

where g_t acts on the frequency $(a2\pi, b2\pi) \subset [0, 2\pi]$ only, with $0 \leq a < b \leq 1$. The best approximation is found by the least square criterion

$$\sum_t \left\{ y_t - \int_a^b \left[\tilde{c}(\omega)\cos(2\pi\omega t) + \tilde{d}(\omega)\sin(2\pi\omega t) \right] d\omega \right\}^2, \tag{8}$$

where minimization is done with respect to $\tilde{c}(\cdot)$ and $\tilde{d}(\cdot)$. This leads to the BP filter

$$\hat{g}_t = \sum_{d:1<t-d} w_{t,d}\, y_{t-d},$$

with weight $w_{t,d}$ given e.g. in Christiano and Fitzgerald (2003). The idea of detrending or smoothing lies in the specification of a and b in (8). In principal a and b are set according to the spectra one wants to extract. Following statistical thinking one might therefore want to choose a and b data driven using according methods. We do not further pursue this idea but formulate the BP filter in an approximate format as penalized spline estimate.

2.4 BP Filter and Penalized Splines

In straight analogy to Section 2.1 we pick up the idea of penalized splines again by employing a basis borrowed from the BP filter above. To do so we replace the integral in (8) by a high but finite dimensional basis. Let therefore

$$B(t) = \big(\cos(2\pi\omega_j t), \sin(2\pi\omega_j t),\ j = 1, \ldots, p\big), \tag{9}$$

be the $2p$ dimensional spline basis with $a = \omega_1 < \ldots < \omega_p = b$ being densely chosen on the interval $[a, b]$. This leads to the summed least square as approximation of (8)

$$\sum_t \{y_t - B(t)\theta\}^2. \tag{10}$$

Instead of simple parametric fitting of θ we impose a penalty on θ. This is also necessary since the dimension $2p$ is chosen large and hence the unpenalized estimates would be wiggled. Wigglyness can thereby be measured by changes in the slope, hence changes in the second order derivative. Note that the second order derivative results through $B(t)\tilde{\theta}$, with

$$\tilde{\theta} = (2\pi)^2 \left(\omega_1^2\theta_1, \omega_1^2\theta_2, \omega_2^2\theta_3, \omega_2^2\theta_4, \ldots\right).$$

We therefore penalize θ_{2j-1} and θ_{2j} with ω_j^2, for $j = 1, 2, \ldots$. To be more specific, let $D = \text{diag}(\omega \otimes 1_2)$ where $\omega = (\omega_1, \omega_2, \ldots)^T$, $1_2 = (1,1)^T$ and $\otimes$ as Kronecker product. This gives the penalized version of (10) as

$$\sum_t \{y_t - B(t)\theta\}^2 + \lambda\, \theta^T D \theta,$$

with λ as penalty parameter as before. Apparently we can also assume a serial correlation for ε_t denoted by R_ε. Also like before we can comprehend the penalty as a priori normal distribution leading to a Mixed Model comparable to (4). The difference lies just in the choice of the basis, but other than that the idea of penalized splines remains the same. This also holds for the estimation of the smoothing parameter.

2.5 Discussion of Proposed Filters

First, both HP and BP filter are grounded on economic arguments to isolate business cycle components from trends. Therefore, the role of the underlying tuning parameters, that is λ in the HP filter and frequencies a and b for the bandpass filter, are set according to the economic interpretation one aims to draw from the business cycle components ε_t. In this respect, no statistical approach or thinking is required. Secondly, both HP and BP filter can be written as minimizer of least squares criteria, i.e. (6) and (8), respectively, does mirror statistical methodology. Third, considering the problem in statistical terms means roughly to seperate Y_t in g_t and ε_t such that ε_t is a stationary process. This task can be carried out with the suggested penalized spline filters using either a basis comparable to the HP or the BP filter, respectively. We will also see that the choice of the basis does not have a relevant influence on the performance of the fit. Comprehending now the business cycle as serial correlation among the ε_t yields a statistical model which can be fitted with penalized splines, as described above, and practically carried out in **R** as shown in the Appendix. The remainder of the paper now demonstrates the four filters in a number of examples with the intention to show that a) the penalized spline approach can compete with HP and BP filter by b) not having any tuning parameter to be chosen by hand or economic grounds.

3 Examples

3.1 Detrending

We now look at practical aspects by considering some of the macroeconomic time series for the United States studied before in Stock and Watson (1999), and described in their Appendix A, see also Harvey and Trimbur (2003). These are quarterly data, with some exception mostly for the time period 1947 to 1996. In order to illustrate the differences in the performance of our filters we are using only a selection of these time series data (with column number referring to Stock & Watson given in brackets), namely real GDP (0), real total consumption (9), real total (fixed) investment (14), employment (27) (which is the average hours worked per employee), real wage rate (44), average labour productivity (33) (which is the output per hour of all persons in nonfarming business), and the yield spread (51), which is the spread between 10 years and 3 months TBill.

For the penalized-spline filter we use a truncated linear basis with $p = 40$ knots. We experimented with a different number of knots as well as a different spline basis (see also below), but no different estimates were observed. This can also be asymptotically justified according to Claeskens, Krivobokova, and Opsomer (2009). Note that this behavior has been demonstrated before in Ruppert (2002). As residual correlation structure in R_ε in (4) we allow for an AR(2) process. We also checked the fitted trend for different stationary correlation structures in the residuals, but the fit remained basically unchanged. This is in line with the results derived in Krivobokova and Kauermann (2007). The HP filter was used in its standard form with $\lambda = 1600$. For the BP filter shown in the right hand side column we set $a = 0.2$ and $b = 1$ (relating to 2 to 30 periods of oscillation) after some experimentation and for the penalized spline fit based on basis (9) we set $p = 40$ and work with an AR(2) process for the serial correlation.

We first look at the log GDP from 1953 to 1996. The corresponding data and the resulting filters are shown in Figure 1. The order of the Figure as well as all subsequent figures is as follows. The top row shows the penalized spline filter as bold curve in comparison with the HP filter in the left panel and in comparison with the BP filter in the right panel.

Apparently both, the HP filter as well as the BP filter tends to mimic parts of the serial correlation structure while the penalized spline filter suggests a simple nearly linear trend. This different specification of the trend is also mirrored in the residuals shown in the middle plot, again compared to residuals based in the HP filter (left column) and the BP filter (right column). The penalized spline filter shows more pronounced residuals with clear serial correlation. The serial correlation can also be seen in the empirical autocorrelation function of the residuals which is shown in the plots in the bottom row of the figure. It is note worthy that the penalized spline filter is data driven and hence adaptive while for the HP filter the smoothing parameter λ it is fixed on economic grounds and so are the tuning parameters in the BP filter.

We proceed with another example and now look at log consumption. The results are shown in Figure 2, following exactly the same ordering of the plots as presented above. The example looks similar to the one above, with the penalized spline fit exhibiting a simple nearly linear trend structure. Looking at the residuals and the autocorrelation, we see that the penalized filters suggest a long range auto-correlation yielding a strong consumption in the residuals in the seventies and late eighties.

As third example we focus on log Investment shown in Figures 3. Here, the BP filter tends to interpolate the data while the HP filter and the penalized spline filters look alike by showing serially correlated residuals.

The fourth example considers log Employment presented in Figures 4. The overall impression remains unchanged as compared to the data examples before. In all three examples, the cyclical components of the spline filter comes out more distinctly compared to the HP and BP filters.

This property also holds for the log Wage data in Figures 5 and the log Productivity data in Figures 6. In all of these graphs again, the penalized spline filter produces a much smoother autocorrelation, which presumably is due to the fact that it allows for serial correlation in the residuals.

Lastly we have added the results of using our three filters for a financial variable. We look here at the trend and cyclical variation of the yield spread, see Figures 7. The yield spread has been the subject of numerous studies in empirical finance and it might be of some interest for the reader how our three filters perform with respect to this variable. As one can observe the trend is linear and slightly upward sloping – an information that might be of interest in the context of a the large number of yield curve studies in empirical finance. Here too the cyclical component for the penalized spline filter comes out more distinct than for the other two filters in particular compared to the BP filter.

We explicitly point out that the above comparison gives an impression of the performance of the penalized spline filter but does not intend to show superiority of any of the three methods. The penalized spline filter, though, is an automatic (or better statistical) routine, as it estimates the appropriate smoothing parameter from the data itself. This resembles some contribution towards objectivity. We also stress that the HP and BP filter fix the smoothing parameter on economic reasons and further tuning would have led to different, possibly more suitable results.

3.2 Cross-Correlation

Another important issue in business cycle analysis has been the extent to which there is a co-movements of time series variables over the business cycles. Technically speaking, the question is thus whether there is some robustness of the filters with respect to a study of the cross-correlations of the residuals. If the GDP variation is, as usual, taken as a standard measure for business cycle fluctuations, then the cross-correlation between the residuals of a macro variable and the GDP, with leads and lags, would give us some information on the co-movement of variables over the business cycle. We look at some cross-correlations resulting from the different filters in order to find out whether the cross-correlation remains robust across different filters. Similar tests on cross-correlation of macro variables with GDP have been undertaken, for example in Cooley (1995), using the HP filter, and Stock and Watson (1999), using the BP filter. Using our above variables, we exemplary look at the cross-correlations for consumption and GDP, investment and GDP, employment and GDP, wage and GDP, labor productivity and GDP, and yield spread and GDP, allowing for four leads and lags for the GDP. Another exercise that we will undertake is to examine the cross-correlation between employment and productivity. This relationship has become central in the Real Business Cycle (RBC) literature, where technology shocks are viewed as driving force for the business cycle. First as seen from Figure 8 (top row) the Spline filter reveals a much more positive cross-correlation with GDP than the HP and BP filters. Since, as numerous macroeconomic studies have shown, consumption is smoothed over the business cycle, the positive correlation with a larger number of leads and lags for the GDP is not unexpected. Up to two leads and lags the results of the three filters show a very similar outcome. Also for the cross-correlation of investment with GDP and employment and GDP the results for the three filters are similar, all three showing positive correlation with GDP up to two leads and lags. Yet the similarity between the penalized spine filter and the HP filter holds also for further leads and lags.

As Figure 9, top row, shows, for the cross-correlation of GDP with the wage all three filters show a very low correlation with GDP, for all four leads and lags. This is a well-known result in macroeconomics that point to the fact that wages are rather sticky over the business cycle. The middle row of Figures 9 depicts the results for the cross-correlation of productivity and GDP. The results shown here again point to the fact that only with two leads and lags the results for the three filters are similar, showing a positive cross-correlation. Yet there are significant non-robust results to be observed for the three filters beyond the two leads and lags.

For our financial variable, the yield spread in the bottom row of Figures 9, there is only a robust result observable for the cross-correlation of output and the yield spread for up to one lead and lag. The positive correlation of the yield spread and GDP, up to one lead and lag, is rather well-known in empirical finance. At high growth rates of the GDP the yield curve usually starts showing an upward slope, since a rise of future interest rates is usually expected. Yet, because of so many other influences on the yield spread, the visible non-robust result for the three filters, beyond one lead and lag, are quite reasonable.

Next we examine the cross correlation between employment and productivity. In the RBC literature, the positive relationship of employment (hours worked) and productivity (GDP per hour) was taken as positive confirmation that technology shocks drive employment. The theoretical model, the RBC model, predicts even a higher correlation of employment with productivity than can be found in the data, since here the productivity is the driving force for output and employment, see Cooley (1995). In Figure 10 we can observe for the penalized spline filter and for the HP filter a strong correlation of employment and productivity for most leads and lags of the two series and a weak correlation between the two series, for leads and lags, for the BP filter. So, overall the data seem to suggest a positive correlation of employment and productivity. We want to note that in the literature sometimes productivity is measured by total factor productivity, the latter is also called Solow residual. Some other researchers use direct productivity (GDP per hour) as measure for productivity as we have done above. In any case recent studies Basu, Fernald, and Kimball (2006) show that either measure of productivity has to be cleansed by eliminating the demand effect on productivity, before it can be used as measure for technology shocks. Thus, productivity increase should be decomposed into the effect that comes from true technology shocks and the effect that comes from demand shocks. This has been done in Basu et al. (2006) whereby the result is obtained that there is no or a weakly negative correlation between employment and productivity.[1] Yet, whatever more detailed studies on the decomposition of productivity in technology and demand shocks will reveal, the overall correlation of employment and productivity is well captured by the penalized spline and HP filter and less so by the BP filter.

4 Conclusion

It appears that the penalized spline filter is indeed a useful alternative to the practically dominating routines like the HP and BP filter. In contrast to these methods the penalized spline filter has two relevant and important advantages. First, it chooses the smoothing or tuning parameter data driven, so subjectivity of the data analyst

[1] Chen et al. (2008) show that although the zero or negative correlation of employment and productivity (using the cleansed productivity measure) seems to hold over the business, in the long run however, the relationship indeed turns out to be positive.

is avoided. Secondly, it allows to decompose a time series into trend and remaining residuals even if there is assumed serial correlation in the residuals. Last but not least, it is numerically handy and implemented for instance in R. The criticism which could be formulated relates to the number of knots being used. Yu and Ruppert (2002) show that the number and location of knots have secondary influence on the performance only. This result is confirmed theoretically and practically following the Mixed Model approach in Kauermann and Opsomer (2009). This includes the case of correlated residuals. Asymptotic investigations how the number of knots should grow with the sample size are given (for independent errors) in Li and Ruppert (2008), Kauermann, Krivobokova, and Fahrmeir (2009) and Claeskens et al. (2009).

The comparison of the performance of our three filters has implications for the recently renewed interest in business cycle research. The HP filter is the preferred filter in Real Business Cycle (RBC) research, see Cooley (1995). The HP filter has recently been complement by the BP filter, see Stock and Watson (1999), where business cycle components are extracted for large number of U.S. macroeconomic times series data. Further far-reaching implications for business cycles theory and causes of recessions, also employing the HP filter, are derived in Chari et al. (2007) and Christiano and Davis (2006). Yet, as our study above shows, those results may need to be qualified, since robust results cannot be obtained across filters, neither with regard to business cycle components nor for the cross-correlation between output and other macro economic times series. The penalized spline has the advantage that it allows for serial correlation.

A Numerical Issues

Function `gamm` in the **R** package `mgcv` fits the data using the Mixed Model representation of penalized splines. This function provides interface with the Mixed Models package `nlme` and allows to estimate the correlation matrix along with the other parameters, as long as the correlation structure is specified. We give here a few lines of a simple **R** code which we used for fitting the data from Section 3. Fitting 200 observations of GDP data with penalized splines, taking an AR(1) correlation structure into account can be carried out as follows.

```
library(mgcv)
data<-scan("GDP.txt")
n=length(data)
t=1:n
k=floor(n/3)
```

```
data.gamm=gamm(data~s(t,k=k,bs="cs"),correlation=corAR1())
plot(t,data)
lines(t,fitted(data.gamm$lme),lwd=3)
```

To fit the data, taking into account an AR(2) correlation structure one has to update the call of `gamm` function as follows.

```
data<-scan("GDP.txt")
data.gamm=gamm(data~s(t,k=k,bs="cs"),correlation=corARMA(p=2))
```

The HP filter and the BP filter are implemented in the `mfilter` package available from the CRAN server of **R**.

References

Basu, S., J. Fernald, and M. Kimball (2006): "Are technology improvements contractionary?" *The American Economic Review*, 96, 1418–1448.

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.

Canova, F. (1998): "Detrending and business cycle facts," *Journal of Monetary Economics*, 41, 475–512.

Canova, F. (2007): *Methods for Applied Macroeconomic Research*, Princeton: Princeton University Press.

Chari, V., P. J. Kehoe, and E. McGrattan (2007): "Busines cycle accounting," *Econometrica*, 75, 781–836.

Chen et al. (2008): "Productivity shocks in the short and long run," *Investigacion Economica*, LXVII, 39–61.

Christiano, L. and J. Davis (2006): "Two flows in business cycle dating," Technical Report 06-12, Federal Reserve Bank of Cleveland.

Christiano, L. J. and T. J. Fitzgerald (2003): "The band pass filter," *International Economic Review*, 44, 435–465.

Claeskens, G., T. Krivobokova, and J. Opsomer (2009): "Asymptotic properties of penalized spline estimators," *Biometrika*, 96, 529–544.

Cogley, T. and J. Nason (1995): "Effects of the Hodrick-Prescott filter on trend and difference stationary time series: Implications for business cycle research," *Journal of Economic Dynamics and Control*, 19, 253278.

Cooley, T., ed. (1995): *Frontiers of Business Cycle Research*, Princeton: Princeton University Press.

Dagum, E. B. and S. Giannerini (2006): "A critical investigation on detrending procedures for non-linear processes," *Journal of Macroeconomics*, 28 (1), 175–191.

Dermoune, A., B. Djehiche, and N. Rahmania (2008): "Consistent estimators of the smoothing parameter in the hodrick-prescott filter," *Journal of the Japan Statistical Society*, 38, 225 – 241.

Dermoune, A., B. Djehiche, and N. Rahmania (2009): "Multivariate extensions of the hodrick-prescott filter-optimality and characterization," *Studies in Nonlinear Dynamics & Econometrics*, 13.

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

Goldrian, G. (2005): "Weaknesses of the Baxter-King filter: Is a pattern-based filter an alternative?" *Jahrbcher fr Nationalkonomie und Statistik*, 255 (4), 386–393.

Harvey, A. and A. Jaeger (1993): "Detrending, stylized facts and the business cycle," *Journal of Applied Econometrics*, 8, 231–247.

Harvey, A. and T. Trimbur (2003): "General model-based filters for extracting cycles and trends in economic time series," *The Review of Economics and Statistics*, 85, 244–255.

Harvey, A., T. Trimbur, and H. Van Dijk (2007): "Trends and cycles in economic time series: A Bayesian approach," *Journal of Econometrics*, 140 (2), 618–649.

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

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.

Kauermann, G., T. Krivobokova, and L. Fahrmeir (2009): "Some asymptotic results on generalized penalized spline smoothing," *Journal of the Royal Statistical Society, Series B*, 71.

Kauermann, G. and J. Opsomer (2009): "Data-driven selection of the spline dimension in penalized spline regression," Submitted to *Journal of the American Statistical Association*.

Kohn, R., C. Ansley, and C. Wong (1992): "Nonparametric spline regression with autoregressive moving average errors," *Biometrika*, 79, 335–346.

Krivobokova, T. and G. Kauermann (2007): "A note on penalized spline smoothing with correlated errors," *Journal of the American Statistical Association*, 102, 1328–1337.

Leser, C. E. V. (1961): "A simple method of trend construction," *Journal of the Royal Statistical Society, Series B*, 23, 91–107.

Li, Y. and D. Ruppert (2008): "On the asymptotics of penalized splines," *Biometrika*, 95, 415–436.

Murray, C. (2003): "Cyclical properties of Baxter-King filtered time," *Review of Economics and Statistics*, 85 (2), 472–476.

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.

Pedregal, D. and P. Young (2001): "Some comments on the use and abuse of the hodrick-prescott filter," *Review on Economic Cycles, International Association of Economic Cycles*, 3.

Proietti, T. (2005): "Forecasting and signal extraction with misspecified models," *Journal of Forecasting*, 24, 539–556.

Ravn, M. O. and H. Uhlig (2002): "On adjusting the Hodrick-Prescott filter for the frequency of observations," *The Review of Economics and Statistics*, 84, 371–380.

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

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

Schall, R. (1991): "Estimation in generalized linear models with random effects," *Biometrika*, 78, 719 – 727.

Schenk-Hopp, K. (2001): "Economic growth and business cycles: A critical comment on detrending time series," *Studies in Nonlinear Dynamics and Econometrics*, 5, 75–86.

Schlicht, E. (2005): "Estimating the smoothing parameter in the so-called Hodrick-Prescott filter," *Journal of the Japanese Statistical Society*, 35, 99–119.

Searle, S., G. Casella, and C. McCulloch (1992): *Variance Components*, Wiley.

Stock, J. H. and M. W. Watson (1999): "Business cycle fluctuations in u.s. macroeconomic time series," in J. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, Amsterdam: Elsevier Science Publishers.

Trimbur, T. (2006): "Detrending economic time series: A Bayesian generalization of the Hodrick-Prescott filter," *Journal of Forecasting*, 25 (4), 247–273.

Wang, Y. (1998): "Mixed effects smoothing spline analysis of variance," *Journal of the Royal Statistical Society, Series B*, 60, 159–174.

Whittaker, E. T. (1923): "On a new method of graduations," *Proceedings of the Edinburgh Mathematical Society*, 41, 63–75.

Wood, S. N. (2000): "Modelling and smoothing parameter estimation with multiple quadratic penalties," *Journal of the Royal Statistical Society, Series B*, 62, 413–428.

Yu, Y. and D. Ruppert (2002): "Penalized spline estimation for partially linear single index models," *Journal of the American Statistical Association.*, 97, 1042–1054.

Zarnowitz, V. and A. Ozyildirim (2005): "Time series decomposition and measurement of business cycles, trends and growth cycles," *Journal of Monetary Economics*, 53, 1717–1739.

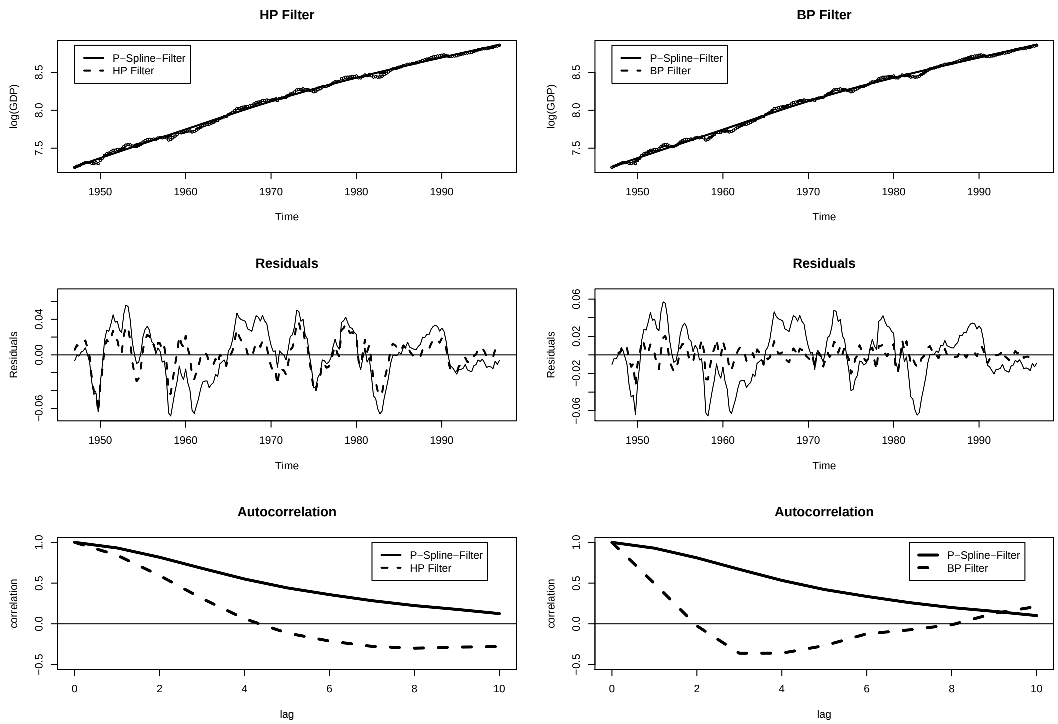

Figure 1: Top row: Log GDP data detrended with penalized splines, taking an AR(2) correlation structure into account. Left column gives comparison to HP filter, right column shows comparison to BP filter.
Middle row: Corresponding residuals.
Bottom row: Autocorrelation for penalized spline residuals and alternatives HP filter (left) and BP filter (right)

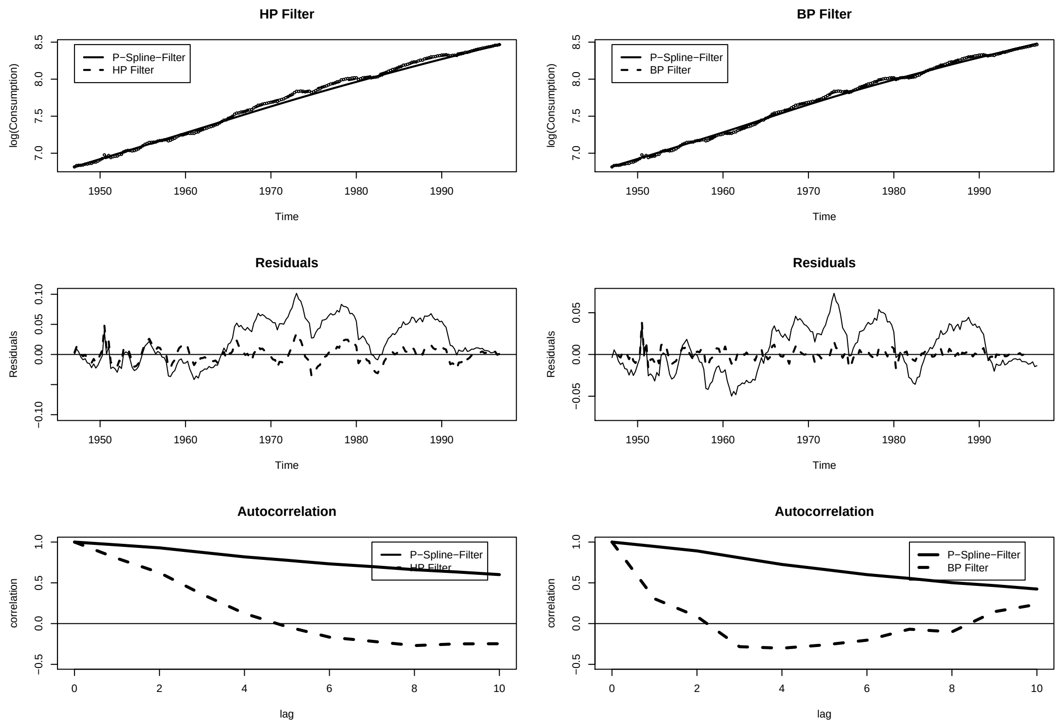

Figure 2: Top row: Log Consumption data detrended with penalized splines, taking an AR(2) correlation structure into account. Left column gives comparison to HP filter, right column shows comparison to BP filter.
Middle row: Corresponding residuals.
Bottom row: Autocorrelation for penalized spline residuals and alternatives HP filter (left) and BP filter (right)

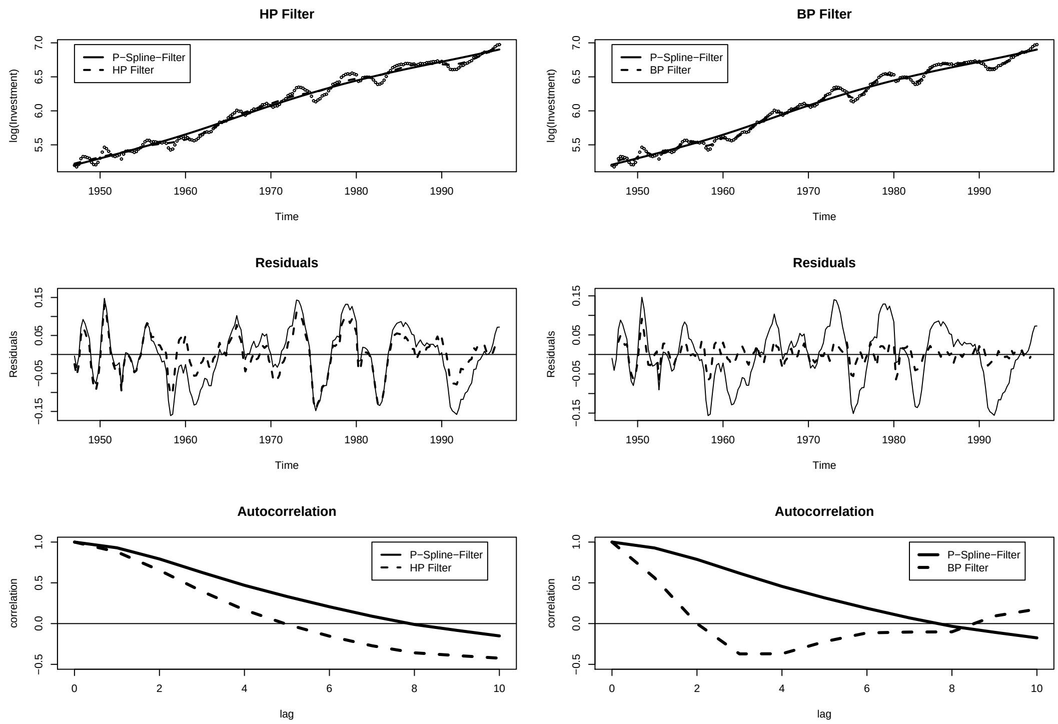

Figure 3: Top row: Log Investment data detrended with penalized splines, taking an AR(2) correlation structure into account. Left column gives comparison to HP filter, right column shows comparison to BP filter.
Middle row: Corresponding residuals.
Bottom row: Autocorrelation for penalized spline residuals and alternatives HP filter (left) and BP filter (right)

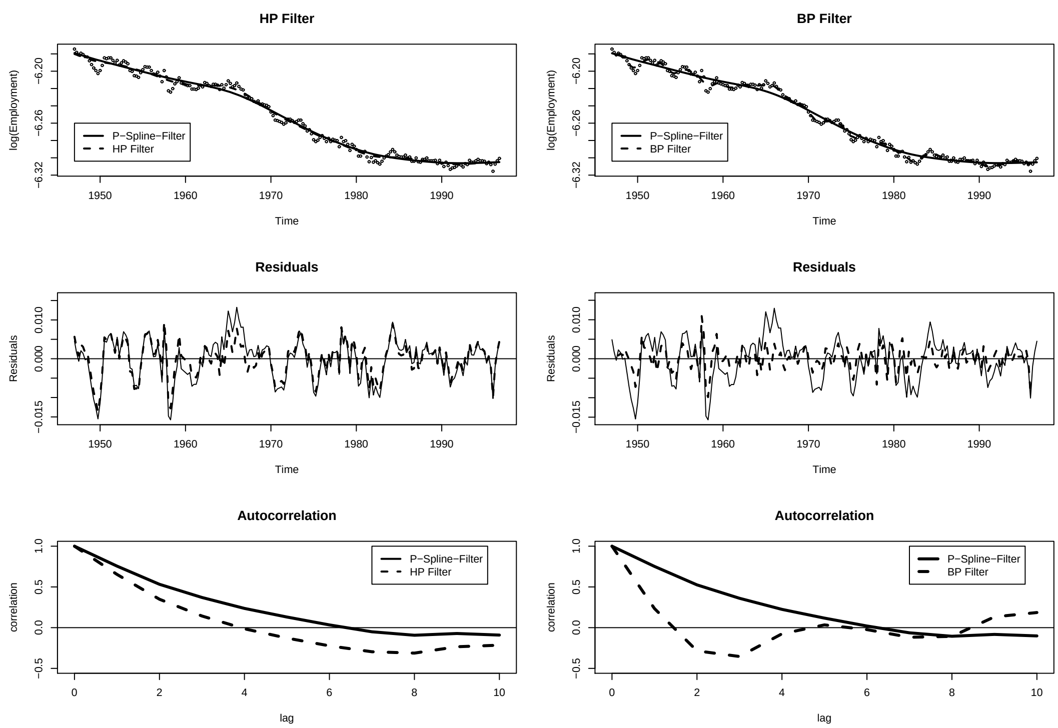

Figure 4: Top row: Log Employment data detrended with penalized splines, taking an AR(2) correlation structure into account. Left column gives comparison to HP filter, right column shows comparison to BP filter.
Middle row: Corresponding residuals.
Bottom row: Autocorrelation for penalized spline residuals and alternatives HP filter (left) and BP filter (right)

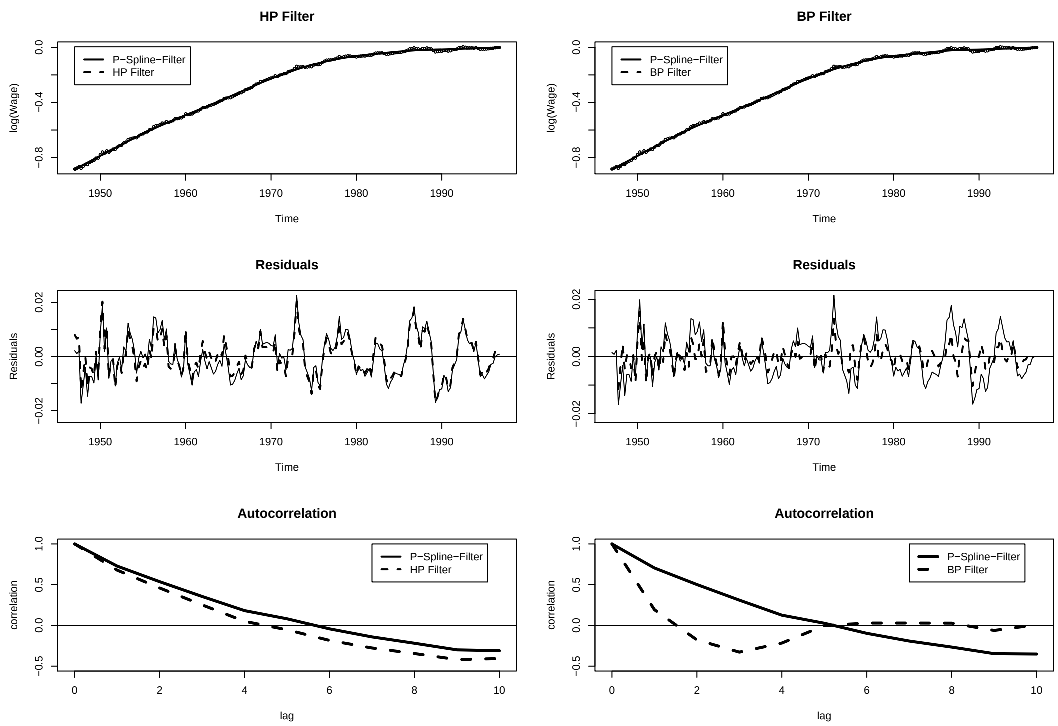

Figure 5: Top row: Log Wage data detrended with penalized splines, taking an AR(2) correlation structure into account. Left column gives comparison to HP filter, right column shows comparison to BP filter.

Middle row: Corresponding residuals.

Bottom row: Autocorrelation for penalized spline residuals and alternatives HP filter (left) and BP filter (right)

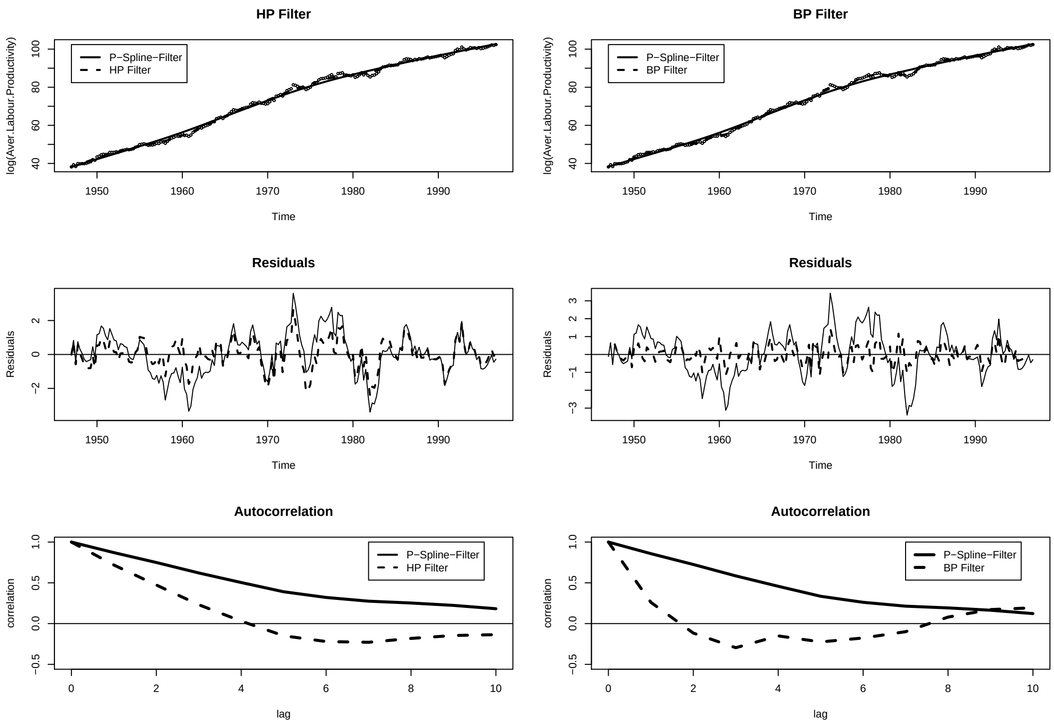

Figure 6: Top row: Log Average Labour Productivity data detrended with penalized splines, taking an AR(2) correlation structure into account. Left column gives comparison to HP filter, right column shows comparison to BP filter.
Middle row: Corresponding residuals.
Bottom row: Autocorrelation for penalized spline residuals and alternatives HP filter (left) and BP filter (right)

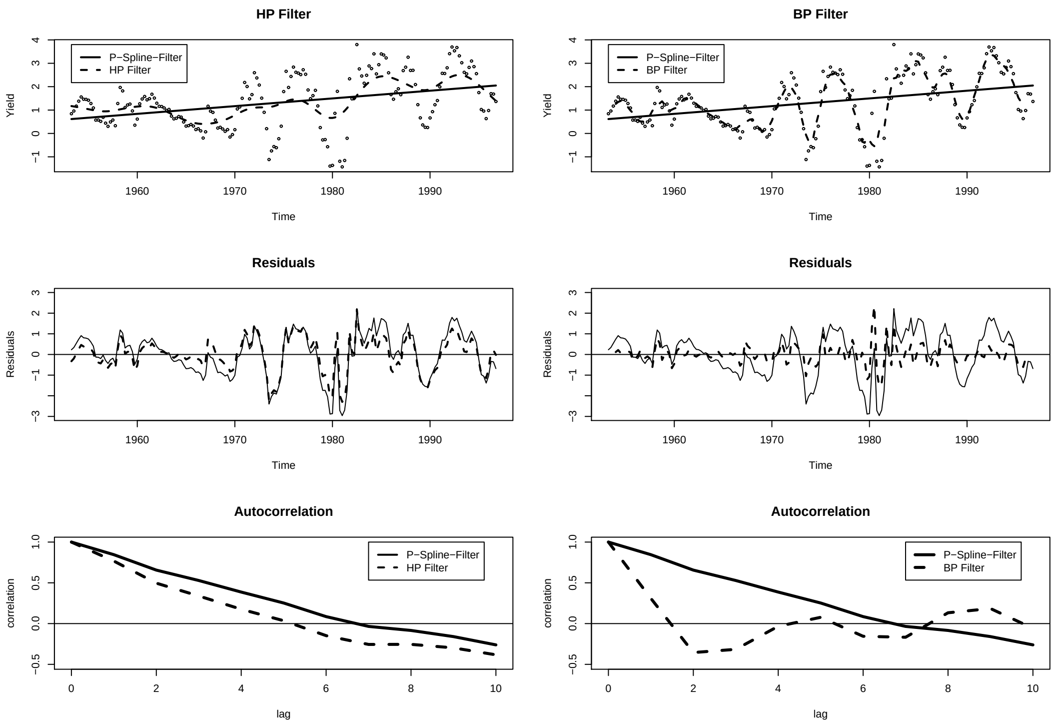

Figure 7: Top row: Yield data detrended with penalized splines, taking an AR(2) correlation structure into account. Left column gives comparison to HP filter, right column shows comparison to BP filter.
Middle row: Corresponding residuals.
Bottom row: Autocorrelation for penalized spline residuals and alternatives HP filter (left) and BP filter (right)

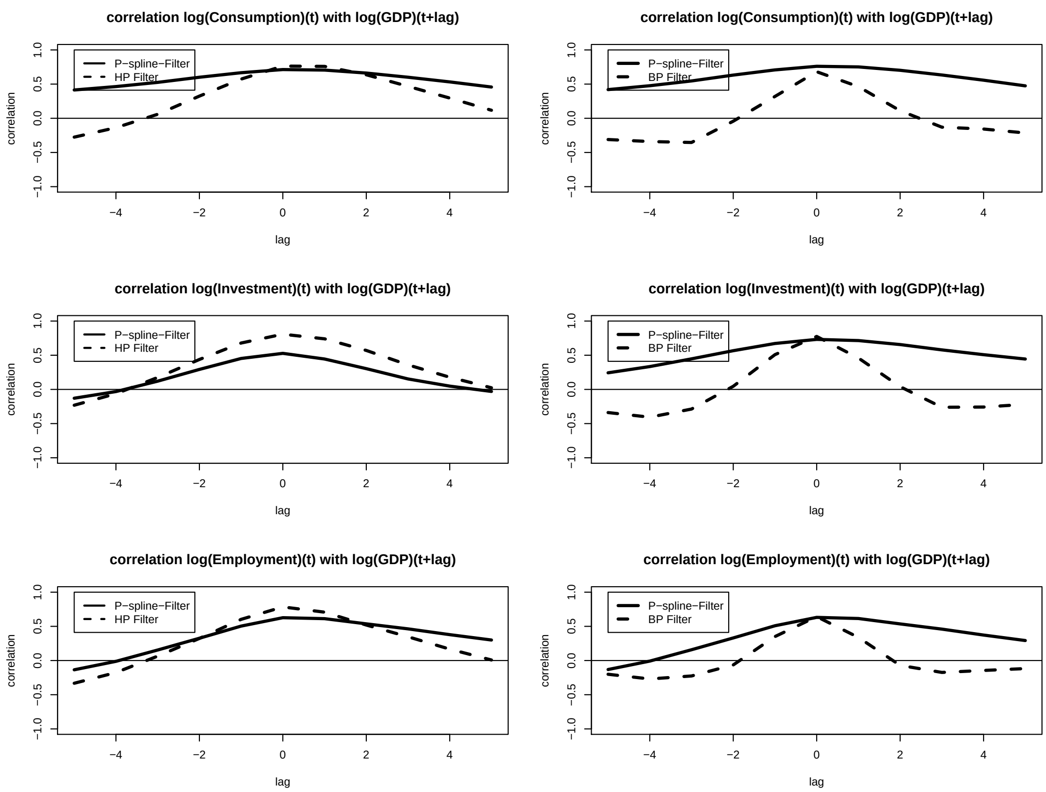

Figure 8: Cross correlation of GDP with Consumption (top row), Investment (middle row) and Employment (bottom row) and comparison with HP filter (left column) and BP filter (right column), respectively

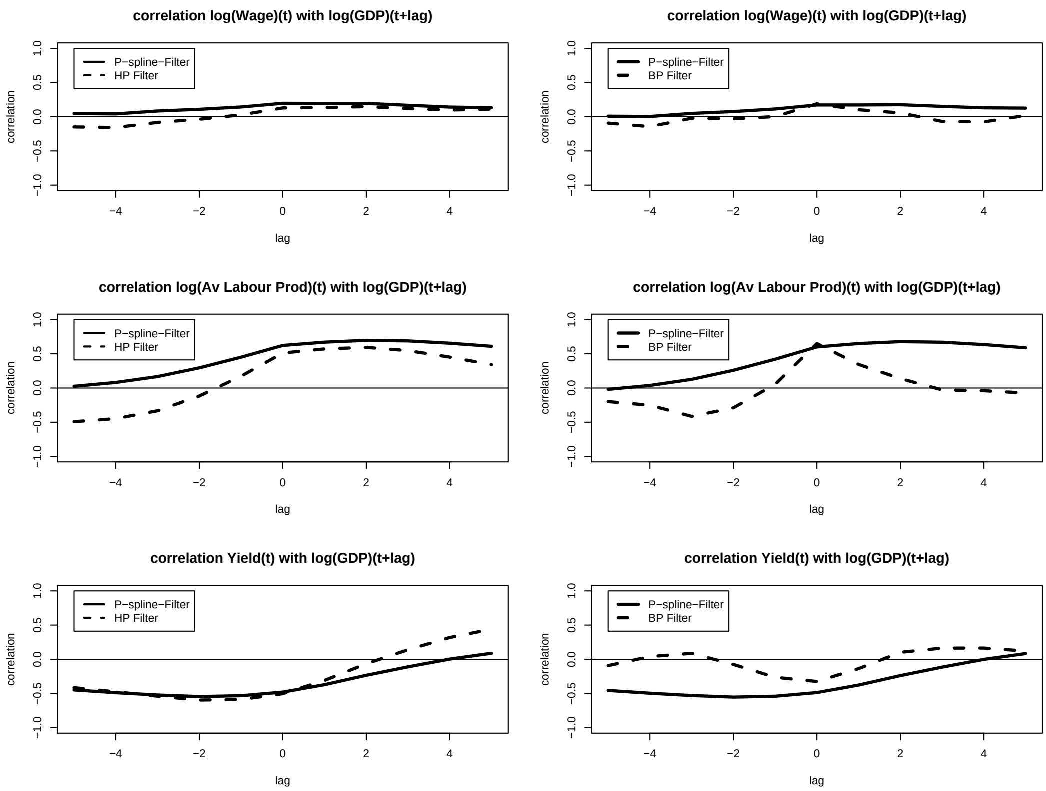

Figure 9: Cross correlation of GDP with Wage (top row), Labour Productivity (middle row) and Yield (bottom row) and comparison with HP filter (left column) and BP filter (right column), respectively

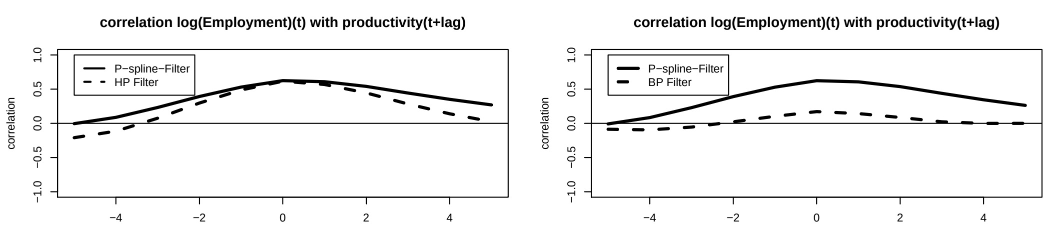

Figure 10: Cross correlation of log Employment and productivity and comparison with HP filter (left) and BP filter (right), respectively