Andreas Blöchl:

Trend Estimation with Penalized Splines as Mixed Models for Series with Structural Breaks

Munich Discussion Paper No. 2014-2

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Trend Estimation with Penalized Splines as Mixed Models for Series with Structural Breaks

Discussion Paper

Andreas Blöchl

February 2014

Abstract

On purpose to extract trend and cycle from a time series many competing techniques have been developed. The probably most prevalent is the Hodrick Prescott filter. However this filter suffers from diverse shortcomings, especially the subjective choice of its penalization parameter. To this point penalized splines within a mixed model framework offer the advantage of a data driven derivation of the penalization parameter. Nevertheless the Hodrick-Prescott filter as well as penalized splines fail to estimate trend and cycle when one deals with times series that contain structural breaks. This paper extends the technique of splines within a mixed model framework to account for break points in the data. It explains how penalized splines as mixed models can be used to avoid distortions caused by breaks and finally provides an empirical application to German data which exhibit structural breaks due to the reunification in 1990.

JEL-Code: C220, C520

Keywords: penalized splines, mixed models, structural breaks, trends, flexible penalization

Andreas Blöchl
Department of Economics
University of Munich
Schackstraße 4
Germany-80539 Munich
andreas.bloechl@lrz.uni-muenchen.de

1 Introduction

An important filed of economic research is the decomposition of a time series into a trend and a cyclical component. The trend represents the long run development of the data, whereas the cycle is supposed to follow any type of stationary process. On purpose to estimate trend and cycle, several different methods have been developed. The probably most popular is the Hodrick-Prescott filter, which traces back to Whittaker (1923), Henderson (1924) and Leser (1961). The Hodrick-Prescott filter (Hodrick/Prescott, 1997) is a linear filter, predominantly depending on the choice of a single penalization parameter λ that controls the smoothness of the estimated trend. The choice of λ has been an issue of research especially since there is no general definition of trend or cycle (Stamfort, 2005) as well as both components are not observable (Mills, 2003). To this point Hodrick/Prescott (1997) suggested a value of $\lambda = 1600$ for quarterly data. Ravn/Uhlig (2002) commended values of 6.25 and 14400 for yearly and monthly data respective.

Nevertheless Hodrick/Prescott's choice of the penalization parameter is often criticized as arbitrary (Danthine/Giradin, 1989), because the derivation of the values can be seen as dubious concerning economic and statistical aspects. The derivation of $\lambda = 1600$ arises from the equivalence of the Hodrick-Prescott filter to the Wiener-Kolmogorov filter (Whittle, 1983; Bell, 1984) when the trend is a twofold integrated random walk and the cycle is white noise as well as subjective assumptions about the variance of trend and cycle (Hodrick/Prescott, 1997). These circumstances are clearly not necessarily given in real time series. Schlicht (2005) showed how the Hodrick-Prescott filter can be written as a linear mixed model in order to estimate the penalization. However, this approach was limited to a white noise error term. Thus, the subjective choice of the penalization parameter is a clear drawback of the Hodrick-Prescott filter.

In contrast to the Hodrick-Prescott filter penalized splines (O'Sullivan, 1986) within a mixed model framework allow to estimate the penalization parameter for a correlated residual structure. Thus, the usually autocorrelated pattern of business cycles can be incorporated into the model. A further advantage of this method is that the results are robust with regard to a misspecification of the correlation structure as long as the misspecification is not too large (Krivobokova/Kauermann, 2007). This is a positive feature as the true correlation structure is unknown in most cases. Paige (2010) shows that the Hodrick-Prescott filter is a special case of a penalized spline. Consequently the mixed model approach with a correlated residual structure is available for this filter. However, the resulting filter is not equivalent to the original Hodrick-Prescott filter any more.

On the estimation of trend and cycle one sometimes faces the problem of structural breaks in the data. Such breaks can be due to many different reasons, for example the contribution of East Germany to the German GDP from 1991 onwards or the abrupt rise in the German population after the reunification. Such break points lead to undesirable distortions, when one is estimating trend and cycle with methods like the Hodrick-Prescott filter or penalized splines. Thus techniques that allow for a consideration of break points are required. Schlicht (2008) suggested to extend the Hodrick-Prescott filter by dummy variables to account for structural breaks. Alternatively Razzak /Richard (1995) and also Pollock (2009) proposed not to use single but different penalization parameters within the Hodrick-Prescott filter. In particular they proposed to change the penalization parameter at the time of the break point to reduce the generated distortions. Even if this method allows accounting for breaks, it still suffers from the shortcoming of the disputed choice of the penalization.

This paper develops a related approach to deal with structural breaks. It extends the mixed model framework for penalized splines to account for breaks by a time varying penalization. This allows to derive a data driven estimation of trend and cycle even if the time series contains a structural break. The paper is structured as follows: The first section shortly summarizes penalized splines with a truncated polynomial basis and how they can be incorporated into a linear mixed model. Afterwards it is shown how the model framework can be extended by a flexible penalization in order to account for breaks points. Finally empirical examples are provided, where especially the method is used to estimate the trend of the German GDP.

2 The model framework

2.1 tp-splines as linear mixed models

This chapter deals with tp-splines. Here tp-splines denote splines with a truncated polynomial basis that trace back to Brumback et al. (1999). This kind of spline is advantageous because of its easy implementation and its obvious connection to linear mixed models. Estimating the trend component with a tp-spline means that the trend function is modeled in dependence of time t. After dividing the exogenous variable t, $t = 1, ..., T$ into $m-1$ intervals by setting knots $1 = \kappa_1 < \kappa_2 < ... < \kappa_m = T$ one can write a tp-spline of degree l for a time series $\{y_t\}_{t=1}^T$

$$y_t = f(t) + \varepsilon_t = \delta_1 + \delta_2 t + \delta_3 t^2 + ... + \delta_{l+1} t^l + \delta_{l+2}(t-\kappa_2)_+^l + ... + \delta_d (t-\kappa_{m-1})_+^l + \varepsilon_t, \quad (1)$$

$$\text{with } (t-\kappa_j)_+^l = \begin{cases} (t-\kappa_j)^l & , t \geq \kappa_j \\ 0 & , \text{else} \end{cases},$$

where $d = m + l - 1$. Writing the model in matrix notation yields:

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}, \quad (2)$$

$$\text{with } \boldsymbol{Z} = \begin{pmatrix} 1 & 1 & \dots & 1^l & (1-\kappa_2)_+^l & \dots & (1-\kappa_{m-1})_+^l \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 1 & T & \dots & T^l & (T-\kappa_2)_+^l & \dots & (T-\kappa_{m-1})_+^l \end{pmatrix}.$$

Here $\boldsymbol{\delta} = (\delta_1, ..., \delta_d)'$, $\boldsymbol{\varepsilon} = (\varepsilon_1, ..., \varepsilon_T)'$ and $\boldsymbol{y} = (y_1, ..., y_T)'$. The truncated polynomials allow a great flexibility of $f(t)$ as they induce that the coefficient of the highest polynomial changes at every knot. The degree of the change and thus the flexibility of $f(t)$ is determined by the values of the coefficients $\delta_{l+2}, ..., \delta_d$. Consequently the flexibility of the trend function can be regulated by controlling the coefficients of the truncated polynomials. This is done by the penalized least squares criterion:

$$PLS(\lambda) = \sum_{t=1}^{T} [y_t - f(t)]^2 + \lambda \sum_{j=l+2}^{d} \delta_j^2. \tag{3}$$

The second part of $PLS(\lambda)$ controls the smoothness of the estimated trend function. It is weighted by a factor λ that is called the penalization parameter. Increasing the value of λ induces a smoother trend as the coefficients of the truncated polynomials become absolutely smaller. Finally the solution of (3) in matrix notation is given to (e.g. Ruppert et al., 2003)

$$\hat{\boldsymbol{\delta}} = (\boldsymbol{Z}'\boldsymbol{Z} + \lambda \boldsymbol{K})^{-1}\boldsymbol{Z}'\boldsymbol{y}, \quad \text{where} \quad \boldsymbol{K} = \text{diag}(\underbrace{0, ..., 0}_{l+1}, \underbrace{1, ..., 1}_{m-2}). \tag{4}$$

$$\text{so that} \quad \hat{\boldsymbol{y}} = \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z} + \lambda \boldsymbol{K})^{-1}\boldsymbol{Z}'\boldsymbol{y}. \tag{5}$$

To receive an estimator for λ a penalized tp-spline can be interpreted as a linear mixed model as already shown by Brumback et al. (1999) (for a detailed discussion of mixed models see Searle et al., 1992; Vonesh/Chinchilli, 1997; Pinheiro/Bates, 2000 or McCulloch/Searle, 2001). A tp-spline within a mixed model has the form (e.g. Ruppert et al., 2003)

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{U\gamma} + \boldsymbol{\varepsilon} = \boldsymbol{Z\theta} + \boldsymbol{\varepsilon}, \tag{6}$$

$$\text{where } \boldsymbol{X} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^l \\ \vdots & \vdots & \ddots & \vdots \\ 1 & T & \dots & T^l \end{pmatrix} \text{ and } \boldsymbol{U} = \begin{pmatrix} (1-\kappa_2)_+^l & \dots & (1-\kappa_{m-1})_+^l \\ \vdots & \ddots & \vdots \\ (T-\kappa_2)_+^l & \dots & (T-\kappa_{m-1})_+^l \end{pmatrix}.$$

Consequently $\boldsymbol{Z} = [\boldsymbol{X}, \boldsymbol{U}]$, $\boldsymbol{\theta}' = [\boldsymbol{\beta}', \boldsymbol{\gamma}']$, $\boldsymbol{\beta} \in \mathbb{R}^{(l+1)\times 1}$ and $\boldsymbol{\gamma} \in \mathbb{R}^{(m-2)\times 1}$. Additional it is assumed that $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{R})$ and $\boldsymbol{\gamma} \sim N(0, \boldsymbol{G})$. This means that a correlated residual structure can be allowed. A special feature of the representation of the spline as a mixed model is its interpretation as a hierarchical model (e.g. Fahrmeir et al., 2009):

$$\boldsymbol{y}|\boldsymbol{\gamma} \sim N(\boldsymbol{X\beta} + \boldsymbol{U\gamma}, \boldsymbol{R}) \text{ and the marginal distribution } \boldsymbol{\gamma} \sim N(0, \boldsymbol{G}).$$

Moreover the spline can be seen as a marginal model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}^* \text{ with } \boldsymbol{\varepsilon}^* = \boldsymbol{U\gamma} + \boldsymbol{\varepsilon}.$$

The distribution of $\boldsymbol{y}$ is then given by $\boldsymbol{y} \sim N(\boldsymbol{X\beta}, \boldsymbol{V})$, where $\boldsymbol{V} = \boldsymbol{R} + \boldsymbol{UGU'}$. Given the marginal model of $\boldsymbol{y}$ the parameter vector $\boldsymbol{\theta}$ can be estimated by maximizing the resulting log-likelihood function with respect to $\boldsymbol{\theta}$ (e.g. Ruppert et al., 2003):

$$\log L(\boldsymbol{\theta}) = -\frac{1}{2}(\boldsymbol{y} - \boldsymbol{Z\theta})'\boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{Z\theta}) - \frac{1}{2}\gamma'\boldsymbol{G}^{-1}\gamma. \tag{7}$$

If it is further specified that $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\Omega}\sigma^2)$ and $\boldsymbol{G} = \text{diag}(\tau^2)$ (e.g. Kauermann et al., 2011), i.e. the error term is correlated with a constant variance σ^2 and the correlation matrix $\boldsymbol{\Omega}$ and the parameters $\boldsymbol{\gamma}$ are uncorrelated and have the constant variance τ^2, then this is equivalent to minimizing (e.g. Fahrmeir et al., 2009)

$$\min_{\boldsymbol{\theta}}\ (\boldsymbol{y} - \mathbf{Z}\boldsymbol{\theta})'\boldsymbol{\Omega}^{-1}(\boldsymbol{y} - \boldsymbol{Z\theta}) + \lambda\boldsymbol{\gamma}'\boldsymbol{\gamma}, \tag{8}$$

where $\lambda = \dfrac{\sigma^2}{\tau^2}$. As $\boldsymbol{G} = \text{diag}(\tau^2)$, the solution of the maximization of formula (7) is (Robinson, 1991; Hayes/Haslett, 1999; also Ruppert et al., 2003)

$$\tilde{\boldsymbol{\theta}} = \begin{bmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\gamma}} \end{bmatrix} = (\boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{Z} + \frac{1}{\tau^2}\boldsymbol{K})^{-1}\boldsymbol{Z}'\boldsymbol{R}^{-1}\boldsymbol{y}, \tag{9}$$

where the estimators $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\gamma}}$ are the best, linear unbiased predictors (BLUPs) of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. A remaining problem is that usually the covariance matrices $\boldsymbol{R}$ and $\boldsymbol{G}$ are unknown and have to be estimated. To this regard both matrices are written in dependence of a vector of parameters $\boldsymbol{\vartheta}$, where $\boldsymbol{\vartheta}$ depends on the assumed correlation structures of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\gamma}$. Then the log-likelihood is derived by the interpretation as a marginal model which is given except of a constant term to (e.g. Ruppert et al., 2003):

$$l(\boldsymbol{\beta}, \boldsymbol{\vartheta}) = -\frac{1}{2}\left[\log(|\boldsymbol{V}(\boldsymbol{\vartheta})|) + (\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{V}(\boldsymbol{\vartheta})^{-1}(\boldsymbol{y} - \boldsymbol{X\beta})\right]. \tag{10}$$

Differentiating and solving with respect to $\boldsymbol{\beta}$ yields

$$\tilde{\boldsymbol{\beta}}(\boldsymbol{\vartheta}) = (\boldsymbol{X}'\boldsymbol{V}(\boldsymbol{\vartheta})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}(\boldsymbol{\vartheta})^{-1}\boldsymbol{y}. \tag{11}$$

Reinserting into (10) finally yields the profile log-likelihood:

$$l_p(\boldsymbol{\vartheta}) = -\frac{1}{2}\left(\log|\boldsymbol{V}(\boldsymbol{\vartheta})| + [\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}(\boldsymbol{\vartheta})]'\boldsymbol{V}(\boldsymbol{\vartheta})^{-1}[\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}(\boldsymbol{\vartheta})]\right). \tag{12}$$

Instead of the profile log-likelihood in most applications the restricted log-likelihood (Searle et al., 1992) is maximized:

$$l_r(\boldsymbol{\vartheta}) = l_p(\boldsymbol{\vartheta}) - \frac{1}{2}\log|\boldsymbol{X}'\boldsymbol{V}(\boldsymbol{\vartheta})\boldsymbol{X}|. \tag{13}$$

The restricted log-likelihood is more accurate in small samples as it takes into account the degrees of freedom for of the fixed effects (Ruppert et al., 2003). By maximizing $l_r(\boldsymbol{\vartheta})$ with respect to $\boldsymbol{\vartheta}$ one receives $\hat{\boldsymbol{\vartheta}}$ and thus also the estimators $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{G}}$. As also σ^2 and τ^2

are contained in $\boldsymbol{\vartheta}$ immediately $\hat{\lambda} = \frac{\hat{\sigma}^2}{\hat{\tau}^2}$ can be received. Furthermore $\hat{\boldsymbol{R}}$ and $\hat{\tau}^2$ can be inserted into (9) which yields the estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ that are called the estimated BLUPs or EBLUPS of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

A further issue is the assumption of the residual correlation structure, as this correlation is usually unknown. To this regard Krivobokova/Kauermann (2007) show that the results are robust, even if the assumed correlation structure is not equal to the real (but unknown) correlation of the cycle. However, the assumed correlation may not be too different from the true one. A further advantage of the interpretation of splines as linear mixed models is that the estimations are almost independent of the number of selected knots as long as it is not too low. Without the interpretation as a mixed model the flexibility of a penalized tp-spline increases with the number of knots. Kauermann/Oppsomer (2011) show that the estimation within a mixed model framework also yields a higher penalization parameter when the number of knots is increased. Thus the estimations are hardly influenced by the number of knots.

2.2 Time varying penalization parameters

To this point it was assumed that the variance of the random effects and thus the penalization is the same at every point in time. This might be reasonable for most economic time series but it can cause distortions, when the series contains structural breaks. A possibility to account for break points is to allow a varying variance of the random effects. A time varying variance was already suggested by Crainiceanu et al. (2004) who modeled the variance as a function that smoothly changes over time. But to avoid distortions by structural breaks, it is not useful to model a smoothly varying penalization parameter. Instead the penalization parameter shall change abrupt at the break points. This can be done as follows:

Assume that there is a structural break at time t^* so that the time series abruptly changes from period t^*-1 to t^*. If equidistant knots are chosen, there will likely be no knots lying exactly on the points in time t^*-1 and t^*. This should be the case if one wants exactly to account for the break in the data, since a function that accurately adapts to the break point is desirable. Therefore knots should be inserted at t^*-1 and t^* in addition to the equidistant knots. Moreover it has to be checked if there is a knot lying between t^*-1 and t^*. In this case either this knot has to be removed or the number of knots m has to be changed. If m equidistant knots are chosen, the number of knots extends to $m+2$, so that the sequence of knots is now given to $\kappa_1, \kappa_2, ..., t^*-1, t^*, ...\kappa_m$. One exception arises if $m = T$, as then there is a knot exactly at the points in time $t = 1, ..., T$. In this case no additional knots have to be inserted.

To achieve a penalization that differs at the break from the one for the rest of the series one allows the variance of the random effects to change at the knots t^*-1 and t^*. Thus the variance of the random effects at the break is now given to υ^2 where the variance is still τ^2 at all other knots. The covariance matrix of the random effects is then given to

$$\boldsymbol{G} = \text{diag}(\underbrace{\tau^2, ..., \tau^2}_{\text{s}-1}, \upsilon^2, \upsilon^2, \underbrace{\tau^2, ..., \tau^2}_{\text{m}-\text{s}-1}). \tag{14}$$

s is the number of knots before $t^* - 1$. Since there are two different variances of the random effects, two penalization parameters are obtained.

$$\lambda_1 = \frac{\sigma^2}{\tau^2} \text{ and } \lambda_2 = \frac{\sigma^2}{\upsilon^2}.$$

λ_2 regulates the penalization at the knots t^* and $t^* - 1$ while λ_1 determines how the trend function can change at all other knots. This model allows to estimate a penalization at the time of the break that is different from the penalization for the rest of the series. Thus the model is able to account for the structural break. Section 3 gives empirical examples and shows that this model indeed yields a penalization that adapts to the break in the data. In order to estimate the tp-spline with the time varying penalization the restricted log likelihood is maximized to receive the estimators $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{G}}$, where $\hat{\boldsymbol{G}} = \text{diag}(\hat{\tau}^2, ..., \hat{\tau}^2, \hat{\upsilon}^2, \hat{\upsilon}^2, \hat{\tau}^2 ..., \hat{\tau}^2)$. Then one defines

$$\hat{\boldsymbol{B}} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \hat{\boldsymbol{G}}^{\boldsymbol{-1}} \end{pmatrix}.$$

Afterwards the parameters can be estimated according to (9) by

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{bmatrix} = (\boldsymbol{Z}'\hat{\boldsymbol{R}}^{-1}\boldsymbol{Z} + \hat{\boldsymbol{B}})^{-1}\boldsymbol{Z}'\hat{\boldsymbol{R}}^{-1}\boldsymbol{y}. \tag{15}$$

2.3 The optimal degree of the tp-spline

One remaining question is what degree for the tp-spline should be chosen. The aim is to construct a trend function, whose estimated slope can change abruptly at a single point in time. The best way to achieve this feature is to use a tp-spline of degree one. As mentioned in chapter 2.1 a tp-spline of degree l is a continuous function that changes the coefficient of the l^{th} polynomial at each knot. With regard to a tp-spline of degree one this means that the estimated trend is just a line that changes its slope at every knot. The change of the slope is regulated by the penalization parameter λ. High values of λ induce that the slope can only change slightly while low values of λ allow large changes of the slope. If the penalization adopts very low values at the break then the slope of the tp-spline of degree one can change abruptly and adapt to the break.

The same is not true for tp-splines of degrees $l \geq 2$. As only the coefficient of the highest polynomial is able to change at each knot, these splines contain coefficients of lower polynomials that do not change, which clearly hinders an abrupt change of the trend function. This is shown in Figure 5, which displays a simulated series that contains a break point.

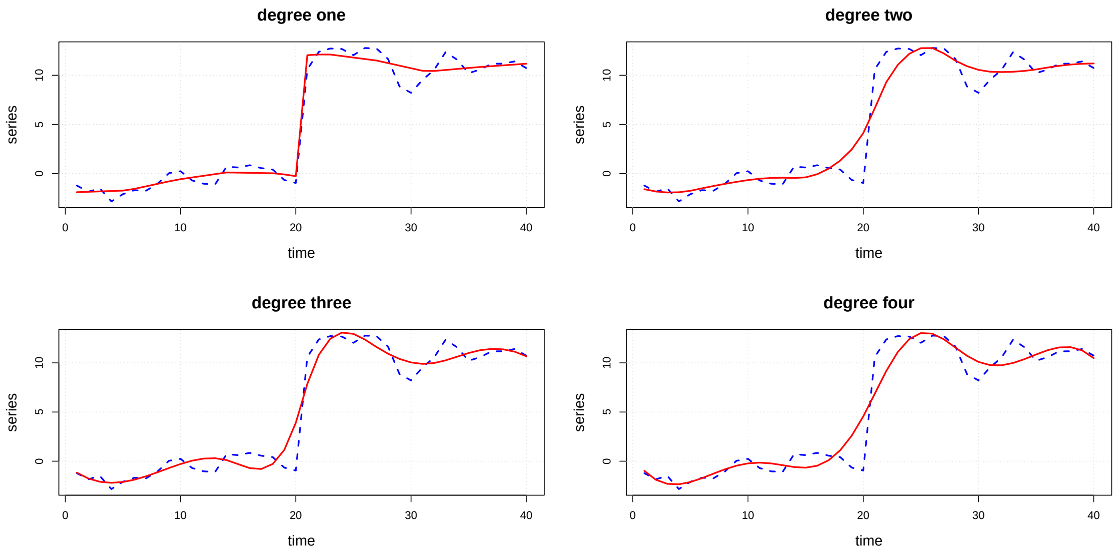

Figure 1: Trend estimation with splines of different degrees

The trend of the series was estimated with penalized tp-splines of degrees one to four using a time varying penalization to account for the break point. One can clearly see that the estimated tp-spline of degree one changes its slope accurately at the break point while all others start rising their slopes too soon and thus are distorted by the break. This is especially true for the spline of degree two, whereas the distortions seems to decrease again, when higher polynomials for the spline are selected.

3 Application to the data

German GDP

In this chapter the described methods are applied to real time series of German data. The German reunification offers good examples for structural breaks since most economic variables change abruptly at the beginning of 1991. To this regard data of the real German GDP from the first quarter 1970 to the second quarter 2013 are considered[1]. The data are adjusted for season and calendar effects.

The analysis of this series is problematic, since it refers only to West Germany before 1991 and contains both East and West Germany afterwards. The German GDP is displayed in Figure 2. One can clearly see a jump in the data from the fourth quarter 1990 to the first 1991. This is due to a change of the base period used for the calculation as well as the reunification and may not be interpreted as a decline of the German GDP.

[1]The data are from the German Federal Statistical Office.

Figure 2: Real German GDP form 1991 to 2013

However this jump causes distortions, when the trend of this series is estimated by a fixed penalization. This is shown in the left plot of Figure 3, which displays the estimated trend of the GDP. The right plot shows the results, when a flexible penalization is used. In both cases the trend was estimated with a tp-spline of degree one, where the number of knots m was set equal to T. Note that this specification is equal to the Hodrick-Prescott filter (Proietti/Luati, 2007; Paige, 2010). If this special spline is estimated within a mixed model assuming that the error term is autocorrelated then it can be seen as an extended version of the HP-filter taking into account a autocorrelated cyclical component. In this case the error term, i.e. the business cycle was assumed to follow an AR(1)-process.

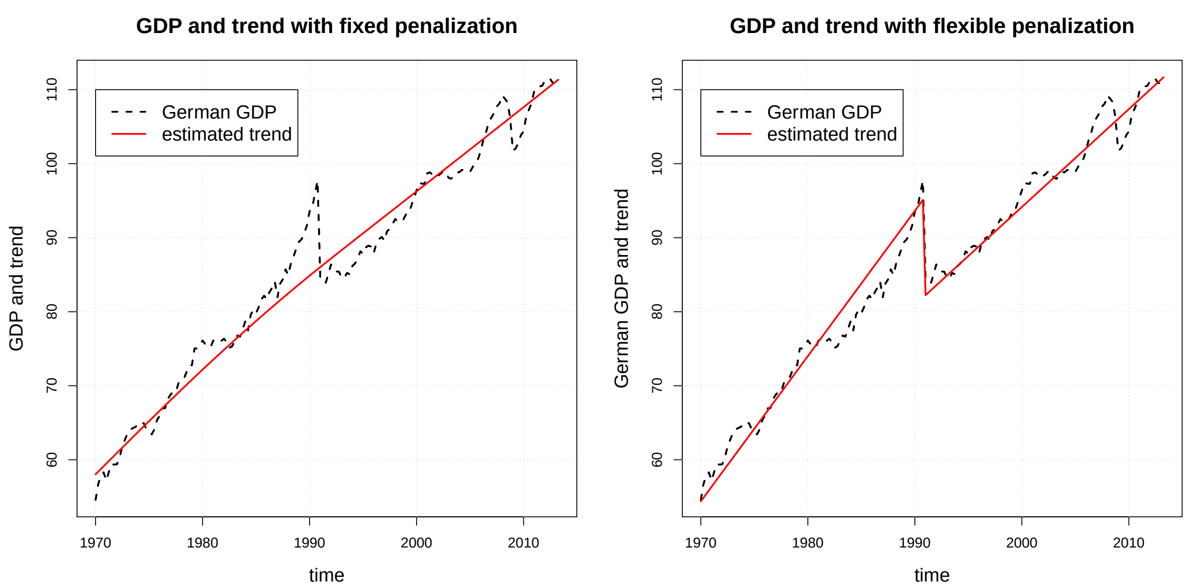

Figure 3: Trend estimation with fixed and flexible penalization and AR(1) residual structure

Clearly, the spline with the fixed penalization cannot account for the break and is thus distorted. The spline with the flexible penalization is able to adapt to the break and can yield a much more reliable estimation for the trend of the German GDP. In both cases the penalization parameter λ was estimated infinitely high, which results in the (almost) linear shape of the trend. Moreover the flexible penalization at the break allowed the estimated trend to abruptly change its slope. This is reasonable as after the reunification also East Germany is included in the GDP. As one can see the slope of the estimated trend has clearly changed after the reunification. While it is 0.49 per quarter before 1991 it is only 0.33 afterwards. Thus the trend growth clearly has declined after the reunification.

To show the effects of another correlation structure of the error term, the trend of the GDP was again estimated assuming that the business cycle follows an AR(2)-process. The results are shown in Figure 4:

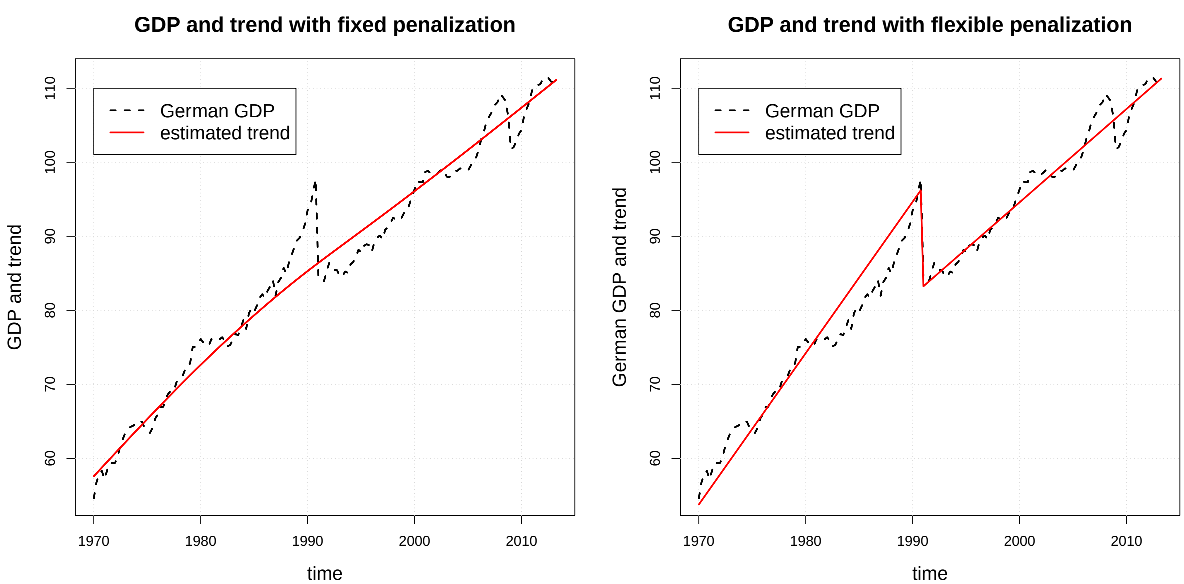

Figure 4: Trend estimation with fixed and flexible penalization and AR(2) residual structure

One can see that there are only slight differences to the case of the AR(1) error tern. There might be a bit more curvature in the trend of the fixed penalization while for the flexible penalization the slope before the reunification is with 0.51 per quarter a bit higher and with 0.32 after the reunification a bit lower than before. Besides the trend component, also the estimated business cycle is affected by the structural break. This is shown in Figure 5, that plots the estimated business cycles for fixed and flexible penalization in the case of an AR(1) error term.

With the exception of the years 2009-2013 the estimated business cycles deviate. Clearly the largest difference is around the years 1990 and 1991, as the cycle according to the flexible penalization excludes the break after the reunification. Moreover, the cyclical component according to the flexible penalization lies significantly above the one of the fixed penalization during the years 1970-1978 and 1991-2008, while it is lower during the period 1978-1991.

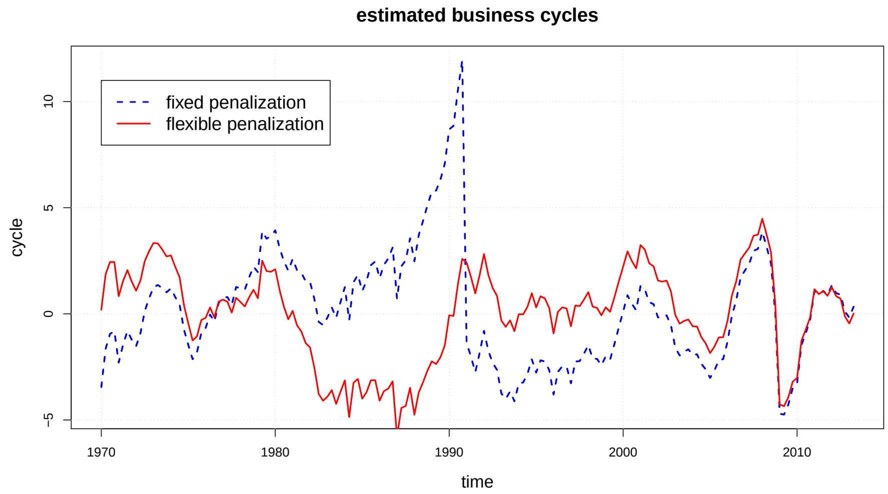

Figure 5: Estimated business cycles for fixed and flexible penalization

This example shows, that it is possible to derive a data driven estimation for trend and cycle of the German GDP even for time series that start before 1991, without having distortions by the reunification. To check for the robustness of the assumed correlation structure finally the autocorrelation- and partial autocorrelation function of the business cycle according to the flexible penalization are considered:

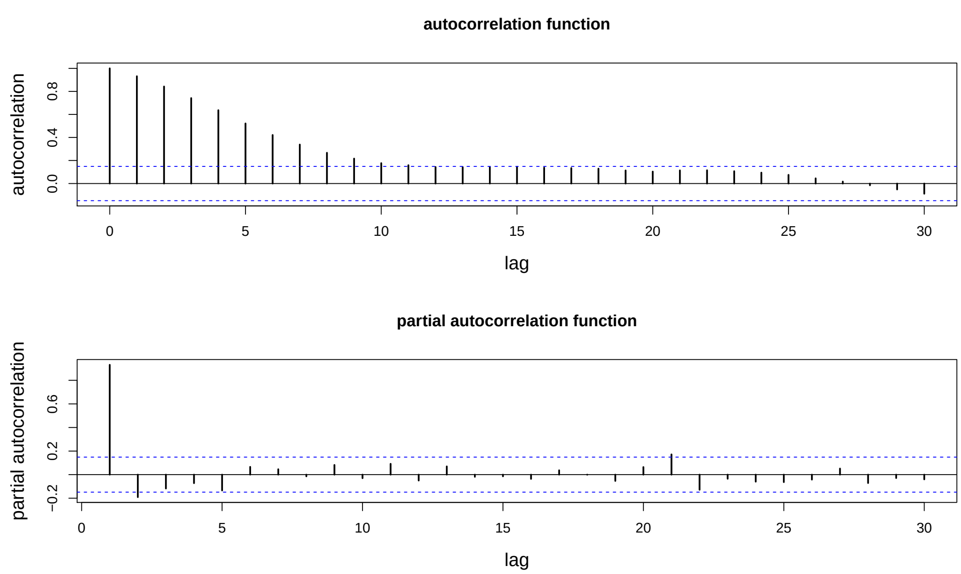

Figure 6: Autocorrelation- amd partial autocorrelation function of estimated cycle

For the estimation within the mixed model framework the cyclical component was assumed to follow an AR(1)-process. The autocorrelation function of the resulting cycle exhibits a damped behavior, while the partial autocorrelation function is only highly significant for the first lag and there is almost no significance after the second lag. Thus the assumption of an cycle that follows an AR(1)-process seems not to be very far away from the resulting correlation structure of the estimated cycle.

German employable population

Another example examined in this paper is the employable population[2] of Germany from the year 1959 to 2010[3]. The data exhibit a jump from the year 1990 to 1991. This is caused by the German reunification, which raised the employable population abruptly from about 43 million to more than 55 million. This time the trend was again estimated using a penalized tp-spline of order one and as many knots as observations, where the error term was assumed to follow an AR(1)-process. Again the estimation was done with a fixed and a flexible penalization. The resulting trend of the fixed penalization is shown on the left side of Figure 7. Because of the jump in the data the estimated trend is just a straight line. This is obviously far away from the real trend since the data suggest an increasing trend before, and a decreasing trend after the reunification.

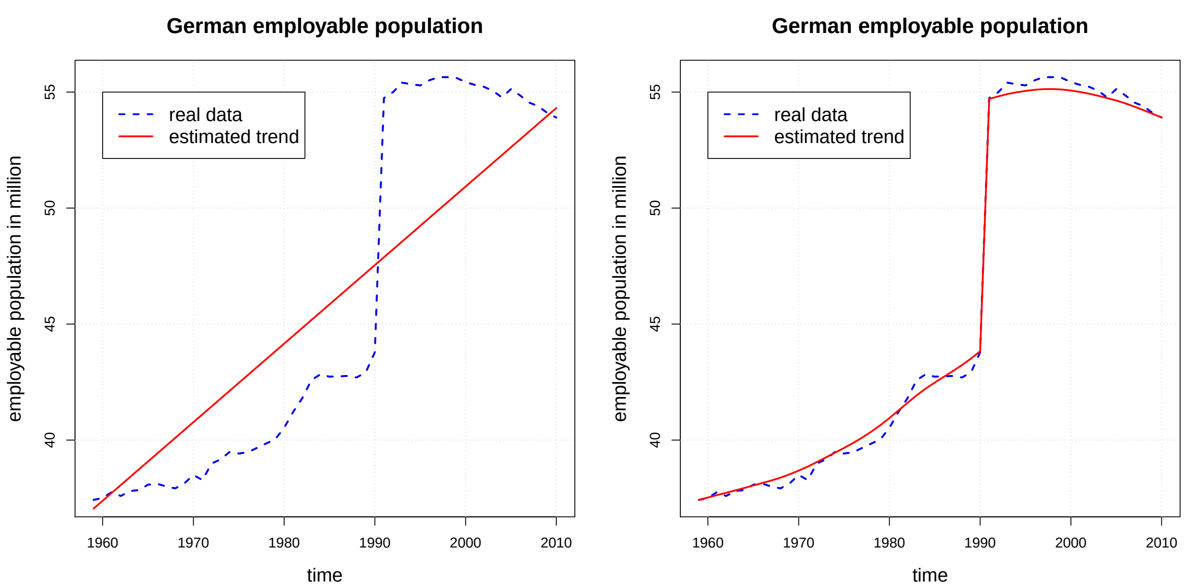

Figure 7: Estimated trend of the employable population with fixed and flexible penalization

The plot on the right side of Figure 7 shows the estimated trend for the flexible penalization, where the penalization was allowed to be flexible at the knots 1990 and 1991. Clearly the estimated trend is not a straight line any more, but it adapts to the break after the reunification. Moreover the trend increases before the year 1990 and decreases afterwards which is in line with the behavior of the expected trend component now.

[2]The employable population is the population of age between 15 and 65.
[3]The data are from the German Federal Statistical Office.

4 Conclusion

This paper shortly summarized penalized splines with a truncated polynomial basis and how they can be incorporated and estimated within a mixed model framework. The interpretation as a linear mixed model allows a data driven derivation of the penalization parameter and helps to overcome the drawback of the subjective choice of the penalization. This approach allows to use a autocorrelated residual structure which is reasonable as the cyclical component is seldom expected to be white noise in economic time series. Beside the penalization parameter penalized splines include more model parameters like the degree of the basis functions, the number of knots and the residual correlation structure. However, in most cases these parameters only slightly influence the resulting trend estimation as shown in Ruppert (2002), Krivobokova/Kauermann (2007), Claeskens et al. (2009) and Kauermann/Oppsomer (2011).

Nevertheless penalized tp-splines within mixed models can fail to estimate trend and cycle, when time series exhibit structural breaks. It was shown in chapter 3 that such breaks can cause distortions in the estimation. Concerning that matter, this paper extended the mixed model framework of penalized splines to account for structural breaks. The variance of the random effects was allowed to change at the break and in the period before. This yielded a time varying penalization that enables the trend function to adapt to the break, where the penalization can still be estimated by maximum likelihood. Moreover it was shown that a tp-spline of degree one is most suitable when one is estimating the trend of time series containing structural breaks. tp-splines of higher degrees are not able to change its slope abruptly enough.

The algorithm described in this paper was demonstrated on empirical examples. Hereby the trend of the real German GDP from 1970-2013 was estimated, where the data exhibit a break at the year 1991 due to the reunification. While the setting with the fixed penalization clearly was distorted downwards by the break, the approach with the flexible penalization could account for this. Given the assumptions of an error term that follows an AR(1)- or an AR(2)-process, the estimation with the flexible penalization yielded an infinitely high penalization parameter and thus a linear trend. Moreover the estimated trend growth rate declined after the reunification.

Thus, based on the ideas of Crainiceanu et al. (2004) who suggested to let the penalization parameter smoothly change over time, and Razzak/Richard (1995) who used a flexible penalization within the Hodrick-Prescott filter to account for breaks, this work offers a useful tool to achieve a data driven estimation of trend and cycle for series that exhibit structural breaks like most German data due to the reunification.

A References

Bell, W. (1984): *"Signal Extraction for nonstationary time series,"* Annals of Statistics, 12, 646-664.

Brumback, B. A., Ruppert, D., Wand, M. P. (1999): Comment on *"Variable selection and function estimation in additive nonparametric regression using a data-based prior,"* Journal of the American Statistical Association, 94, 794-797.

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

Crainiceanu, C., Ruppert, D., Carroll, R., (2004): *"Spatially adaptive Bayesian P-splines with heteroscedastic errors."* Johns Hopkins University Dept. of Biostatistics Working Paper Series No. 1061, Berkeley Electronic Press.

Danthine, J., Girardin, M., (1989); *"Business Cycles in Switzerland. A Comparative Study,"* European Economic Review, 33(1), S. 31-50.

Fahrmeir, L., Kneib, T., Lang, S., (2009): *"Regression - Modelle, Methoden und Anwendungen,"* Springer Verlag, Berlin Heidelberg.

Hayes, K., Haslett, J. (1999): *"Simplifying general least squares,"* American Statistician, 53, 376-381.

Henderson, R. (1924): *"On a new method of graduation,"* Transactions of the Actuarial Society of America, 25, 29-40.

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

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

Kauermann, G., Krivobokova, T., Semmler, W., (2011): *"Filtering Time Series with Penalized Splines,"* Studies in Nonlinear Dynamics & Econometrics, 15(2), Article 2.

Kauermann, G., Opsomer, J., (2011): : *"Data-driven selection of the spline dimension in penalized spline regression,"* Biometrika, 98(1), 225-230.

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

McCulloch, C. E., Searle, S. R. (2001): *"Generalized, Linear, and Mixed Models,"* New York: Wiley.

Mills, T.C. (2003): *"Modelling Trends and Cycles in Economic Time Series,"* Palgrave Macmillan. Houndmills/Basingstoke/Hampshire/New York.

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

Paige, R. L. (2010): *"The Hodrick-Prescott filter: A special case of penalized spline smoothing,"* Electronic Journal of Statistics, 4, 856-874.

Pinheiro, J. C., Bates, D. M. (2000): *"Mixed-Effects Models in S and S-PLUS,"* New York: Springer-Verlag.

Pollock D. S. G. (2009): *"Investigating economic trends and cycles,"* Boumans, M. (Ed.), Palgrave Handbook of Econometrics, Volume 2: Applied Econometrics, Palgrave Macmillan, 243-307.

Proietti, T., Luati, A. (2007): *"Least Squares Regression: Graduation and Filters,"* Boumans, M. (Ed.), Measurement in Economics: A Handbook, ch. 16, Academic Press.

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

Razzak, W., Richard, D., (1995): *"Estimates of New Zealand's Output Gap using the Hodrick Prescott filter with an non-constant Smoothing Parameter,"* Reserve Bank of New Zealand, Discussion Paper G95/8.

Robinson, G. K. (1991): *"That BLUP s a good thing: The estimation of random effects,"* Statistical Science, 6, 15-51.

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

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

Schlicht, E. (2005): *"Estimating the Smoothing Parameter in the so-called Hodrick-Prescott Filter,"* Journal of the Japanese Statistical Society, Vol. 35 No. 1, 99-119.

Schlicht, E. (2008): *"Trend Extraction from Time Series with Structural Breaks and Missing Observations,"* Journal of the Japanese Statistical Society, Vol. 38 No. 2, 205-292.

Searle, S. R., Casella, G., McCulloch, C. E. (1992): *"Variance Components,"* New York: Wiley.

Stamfort, S., (2005): *"Berechnung trendbereinigter Indikatoren für Deutschland mit Hilfe von Filterverfahren,"* Deutsche Bundesbank, Discussion Paper Nr. 19.

Vonesh, E. F., Chinchilli, V. M. (1997): *"Linear and Nonlinear Models for the Analysis of Repeated Measures,"* New York: Marcel Dekker.

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

Whittle, P. (1983): *"Prediction and Regulation,"* Blackwell.