Andreas Bloechl:

Reducing the Excess Variability of the Hodrick-Prescott Filter by Flexible Penalization

Munich Discussion Paper No. 2014-1

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Reducing the Excess Variability of the Hodrick-Prescott Filter by Flexible Penalization

Working Paper

Andreas Blöchl

January 2014

Abstract

The Hodrick-Prescott filter is the probably most popular tool for trend estimation in economics. Compared to other frequently used methods like the Baxter-King filter it allows to estimate the trend for the most recent periods of a time series. However, the Hodrick-Prescott filter suffers from an increasing excess variability at the margins of the series inducing a too flexible trend function at the margins compared to the middle. This paper will tackle this problem using spectral analysis and a flexible penalization. It will show that the excess variability can be reduced immensely by a flexible penalization, while the gain function for the middle of the time series is used as a measure to determine the degree of the flexible penalization.

JEL-Code: C220, C520.

Keywords: Hodrick-Prescott filter, spectral analysis, trend estimation, gain function, flexible penalization

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

1 Introduction

The Hodrick-Prescott filter (Hodrick/Prescott, 1997) is one of the most popular tools for trend estimation in economics. Its advantages are clearly an easy and numerical fast and stable implementation, while the shape of the estimated trend completely depends on the choice of a single penalization parameter λ. In most applications λ is set to 1600 for quarterly data according to the suggestion of Hodrick/Prescott (1997), which can be seen as an "industrial standard" (Flaig, 2012) for economic trend estimation. However this choice is often criticized in literature as dubious (Danthine/Giradin, 1993), not data driven (e.g. Schlicht, 2005; Kauermann et al., 2011) and too low for most time series (Mc Callum, 2000; Flaig, 2012). As an alternative Schlicht (2005) described how the Hodrick-Prescott filter can be incorporated into a mixed model framework in order to derive a data driven estimation of λ. This approach is based on the assumption of a white noise business cycle, but it was shown that this approach can be extended to account for correlated residuals (e.g. Proietti, 2007; Proietti/Luati, 2007). Another possibility to choose λ is suggested by Flaig (2012), who proposed to select λ such high, that the trend component doesn't feature any cyclical behaviour any more.
Nevertheless the choice of $\lambda = 1600$ can be justified by spectral analysis. Using quarterly data this selection implies that the trend approximately consists of oscillations with periodicities of above nine to ten years (Tödter, 2002; Maravall/del Rio, 2001), which is reasonable according to economic conceptions of trend and cycle (e.g. Baxter/King, 1999).

Especially if the selection of λ is based on frequency domain aspects, a further arising problem of the Hodrick-Prescott filter is the increasing excess variability at the margins of a time series. A rising excess variability means that the filter cannot suppress high frequencies for the first and last few estimations any more. This induces that the volatility at the margins of the estimated trend increases compared to the rest of the series, making the trend estimation for the first and last periods more flexible than desired. This is a drawback as researchers and politicians are mainly interested in the trend of the most recent periods that is deterred by the excess variability.

An often applied approach to solve the problem of the excess variability is to use ARIMA models. ARIMA models are used to derive forecasts, that are attached to the time series. This way the original margin of the series moves closer to the middle of the data and is thus less affected by the excess variability. However, this method exhibits the drawback that the forecasts feature failures, that rise with an increasing forecast horizon. Consequently also the estimated trend is subject to this uncertainty.

This paper will tackle the problem of the excess variability using a flexible penalization and spectral analysis. A changing penalization was already introduced for the HP-filter by Razzak/Richard (1995) in order to account for structural breaks and by Crainiceanu et al. (2005) for penalized splines (O'Sullivan, 1986) within an mixed model framework. It is shown that the volatility at the margins can be reduced by letting λ increase to the margins of the series. In order to determine the degree of the rise of the penalization, the gain function of the Hodrick-Prescott filter in the middle of the data is taken as a measure. The flexible penalization is selected such that the gain function for the estimations at the margins is adjusted to the one in the middle.

This paper will not discuss the selection of λ in general. Instead it will stick to the standard choice of $\lambda = 1600$ and show how the excess variability can be reduced for this selection. The first section of this paper will briefly review the Hodrick-Prescott filter. Afterwards the features of the filter in the frequency domain will be examined, especially its characteristics at the margins. Then it will be shown how the gain function can be used as a measure for the excess variability and how flexible penalization can reduce the increasing volatility at the margins. Finally the paper will give some empirical examples and point out the different implications between the results of the flexible penalization and the standard approach.

2 The Hodrick-Prescott filter

2.1 General framework

The Hodrick-Prescott filter (HP-filter) decomposes a time series $\{y_t\}_{t=1}^{T}$ in two components

$$y_t = \mu_t + c_t, \tag{1}$$

where μ_t is the trend and c_t is the business cycle. μ_t is estimated by solving te following minimization problem:

$$\min_{\mu_t} \sum_{t=1}^{T} (y_t - \mu_t)^2 + \lambda \sum_{t=2}^{T-1} [(\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1})]^2. \tag{2}$$

The minimization problem consists of two parts. The first one is the squared deviation of the trend from the original series. Minimizing the first part only yields a trend that is identical to $\{y_t\}_{t=1}^{T}$. The second part consists of the squared second differences and is a measure for the volatility of the trend. Minimizing the second part only would yield a linear trend. Thus there is clearly a trade off between both parts. This trade off is solved by the so called penalization parameter λ that puts weight on the second part. Thus the smoothness of the estimated trend can be completely regulated by the selection of λ. A low value of λ will generate a flexible trend whereas high values make the trend become linear (Stamfort 2005). While the positive features of the HP-filter are clearly its easy and fast numerical implementation as well as its ability to derive estimations at the margins of the series, it suffers from shortcomings like phase shifts (King/Rebelo 2003), spurious cycles (Cogley/Nason 1995) and that there are no general rules for the selection of λ.

Solving the minimization problem in (2) yields the solution of the filter in matrix notation (Mc Elroy 2008):

$$\hat{\boldsymbol{\mu}} = (\mathbf{I} - \lambda \boldsymbol{\Delta}' \boldsymbol{\Delta})^{-1} \mathbf{y}, \tag{3}$$

with $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, ..., \hat{\mu}_T)'$ and $\mathbf{y} = (y_1, ..., y_T)'$. $\boldsymbol{\Delta}$ is a $(T-2) \times T$ differencing matrix,

$$\boldsymbol{\Delta} = \begin{pmatrix} 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 \end{pmatrix},$$

where the product of $\boldsymbol{\Delta}$ and $\mathbf{y}$ yields the second differences of $\mathbf{y}$.

2.2 The Hodrick-Prescott filter in the frequency domain

It is useful to consider the HP-filter in the frequency domain, as its gain function will be applied as a measure to describe the excess variability at the margins. To this point, first of all the structure of the filter weights is considered. Given formula (3) the filter weights are contained in the matrix $(\boldsymbol{I} - \lambda\boldsymbol{\Delta}'\boldsymbol{\Delta})^{-1} \hat{=} \boldsymbol{H} \in \mathbb{R}^{T\times T}$, where the t^{th} row of $\boldsymbol{H}$ contains the weights for the estimation $\hat{\mu}_t$, i.e.

$$\hat{\mu_t} = \sum_{j=1}^{T} h_{tj} y_j. \tag{4}$$

h_{tj} is the j^{th} element of the t^{th} row of $\boldsymbol{H}$. An important feature of this weight matrix is, that the filter weights have a very similar, almost symmetric structure for estimations in the middle of the data, while this structure changes to the margins. This can be seen in Figure 1 that plots the weights for different estimations for a HP-filter with $\lambda = 1600$ that is applied to a series with 100 observations.

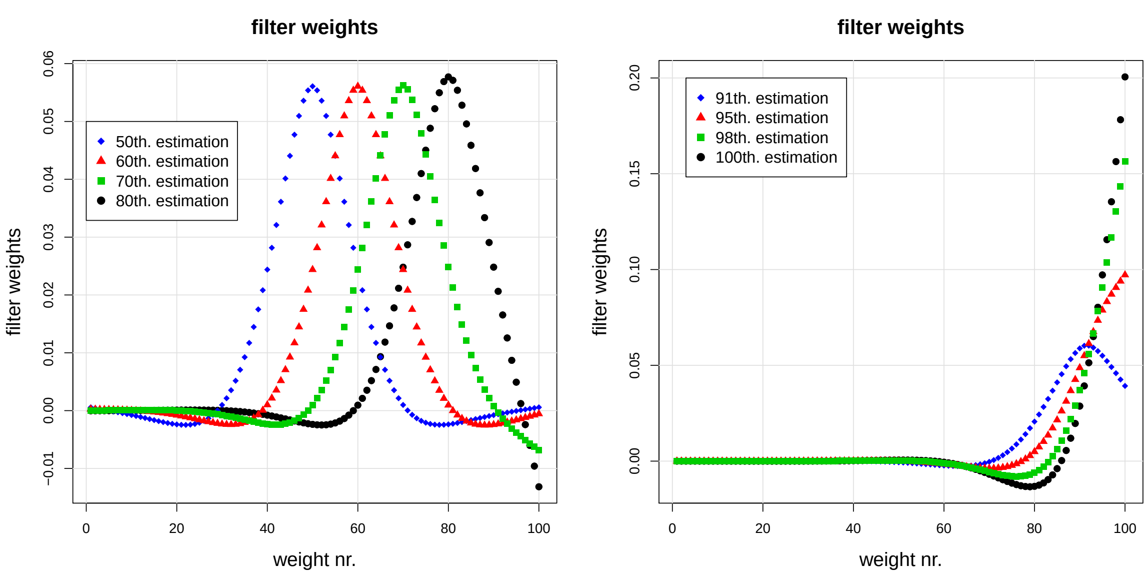


Figure 1: Filter weights in the middle and at the margin

Clearly the filter weights in the middle look similar and converge symmetrically to zero. To the margins the filter weights become more and more asymmetric. Furthermore the highest weight of the estimations at the margins is far above of those around the middle. That means that the estimations at the margins are much more affected by single observations than the estimations closer to the middle of the data. This behaviour of the filter weight structure causes the so called excess variability at the margins, as a single observation can heavily influence the estimated trend.

The change in the weight structure to the margins also becomes obvious, when the HP-filter is considered in the frequency domain. In the frequency domain a time series is interpreted as the overlap of oscillations with different frequencies (for a detailed discussion see Harvey (1993), Hamilton (1994) or Mills (2003)), where the trend as the long run development of a series is supposed to consist of those oscillations with a high periodicity. The HP-filter as a tool for trend estimation extracts oscillations with high periodicities and eliminates oscillations with lower periodicities. This behaviour can be described by the gain function. Given the filter weights h_{tj} that arise for a certain value of λ, the gain for estimation $\hat{\mu}_t$ and frequency ω can be calculated as (e.g. Mills, 2003)

$$g_t(\omega, \lambda) = \sqrt{\left(\sum_{j=1-t}^{T-t} h_{t,j+t} \cos(\omega j)\right)^2 + \left(\sum_{j=1-t}^{T-t} h_{t,j+t} \sin(\omega j)\right)^2}\,. \tag{5}$$

The gain is interpreted as the factor by which the amplitude of an oscillation with a certain frequency is damped or amplified by a filter. In order to show the effects of the changing filter weight structure, Figure 2 displays the gain functions for different estimations.

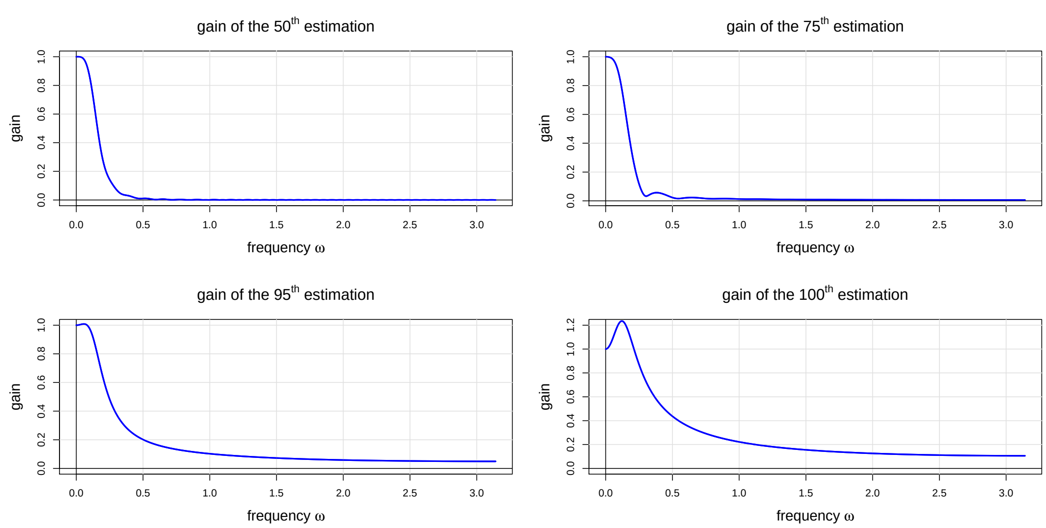


Figure 2: Gain functions for different estimations

For the 50^{th} and 75^{th} estimation the gain functions are very similar as they are induced by an almost equal weight structure. However, as one approaches the margins of the series, the gain function starts to change like for the 95^{th} and 100^{th} estimation. For the estimations at the margins high frequencies can not be completely eliminated any more, which causes an increasing volatility of the resulting trend function. This increase of the volatility is known as the excess variability. It is not practicable to consider the gain functions for all estimations of the time series. Thus a measure is desirable that allows to easily describe and quantify the rise of the volatility in the trend estimation. Such a measure can be the deviation of the gain function of a certain estimation $\hat{\mu}_t$ from the one of the medium estimation. Let $g_m(\omega, \lambda)$ denote the gain for frequency ω for the medium estimation $\hat{\mu}_m$, where $m = \text{ceiling}(T/2)$, and $g_t(\omega, \lambda)$ the one for estimation $\hat{\mu}_t$, then a *loss* can be defined:

$$l(t, \lambda) = \int_0^{\pi} [g_m(\omega, \lambda) - g_t(\omega, \lambda)]^2 d\omega. \tag{6}$$

$l(t, \lambda)$ is the squared deviation of the gain for estimation $\hat{\mu}_t$ from the one for the medium estimation in the interval $[0, \pi]$. The calculation of (6) for continuous frequencies is difficult, however it can be easily approximated by a sufficient high number of discrete frequencies, e.g. for $\boldsymbol{\omega} = (0, 0.001, 0.002, ..., \pi)'$:

$$l(t, \lambda) = \sum_{i=1}^{n} [g_m(\omega_i, \lambda) - g_t(\omega_i, \lambda)]^2 \cdot \delta. \tag{7}$$

n is the number of elements in $\boldsymbol{\omega}$ and δ is the distance between the elements in $\boldsymbol{\omega}$, i.e. $\delta = \omega_j - \omega_{j-1}$. Calculating the *loss* for all $t = 1, ..., T$ gives an overview of which estimations are affected by the increase of the variability. An important question is which factors influence the excess variability. To shed light on this issue first of all a HP-filter with $\lambda = 1600$ is applied to time series of length 50, 100, 150 and 200. Then the *loss* is calculated for each element of these series, which is defined as the *loss function*.

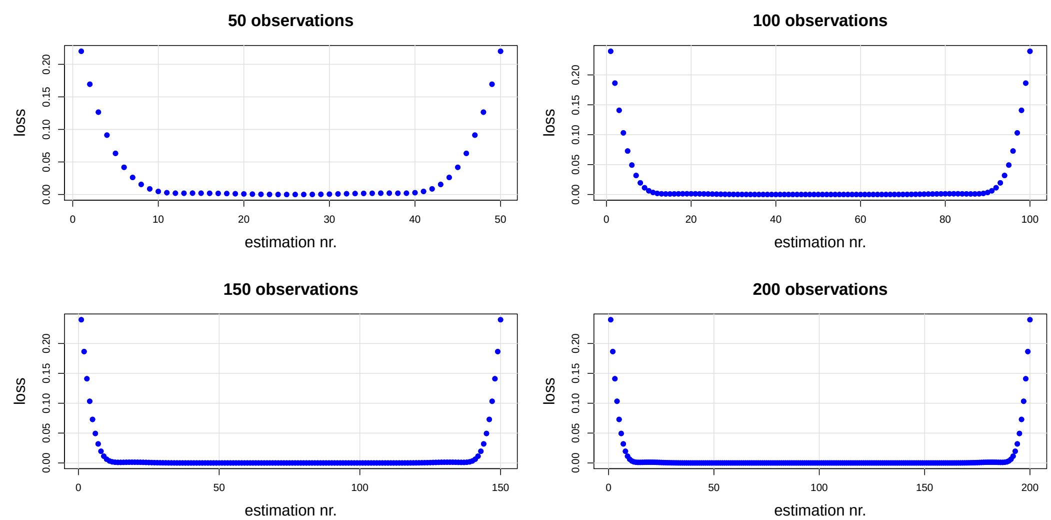


Figure 3: *loss function* for varying data length

As Figure 3 shows, the *loss* is very similar and almost zero for the estimations in the middle of the data and starts to increase abruptly at the margins, which is in line with Figures 1 and 2 that indicate, that the gain of the HP-filter only changes for the first and last few estimations. Independent of the data length, the first and last ten or eleven estimations exhibit a rise of the *loss*. Consequently the excess variability does not depend on the number of observations.

However it can be shown that the number of affected estimations depends on the value of λ. To this regard Figure 4 displays the *loss* for HP-filters with different values of λ, that are applied to a series with 100 observations:

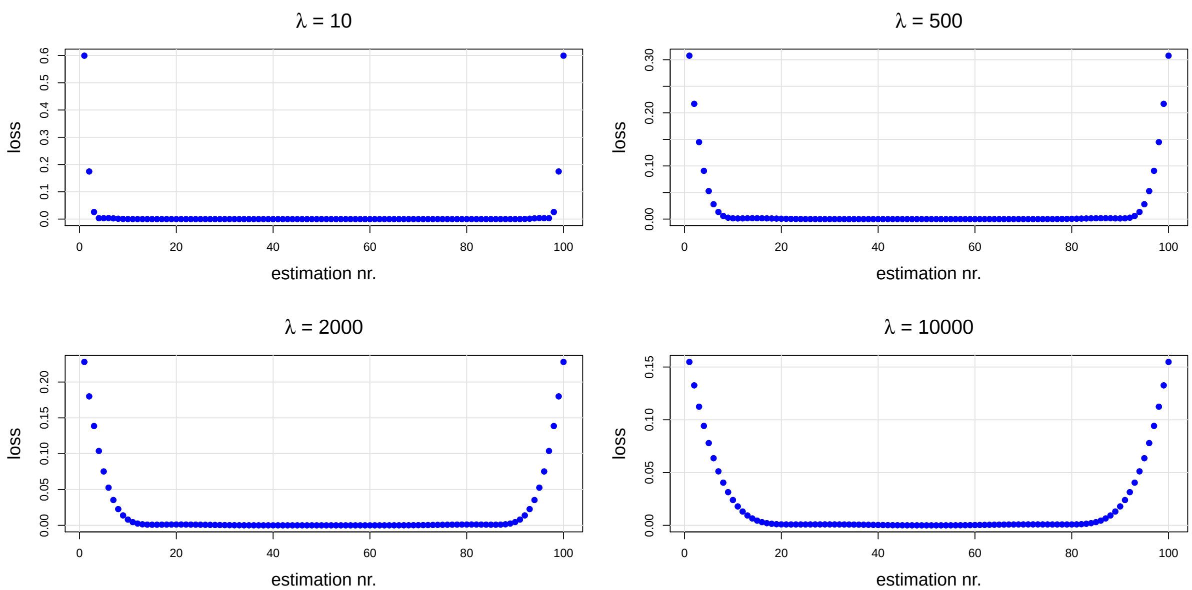


Figure 4: $loss\ function$ for different values of λ

In contrast to the length of the series, the value of λ affects the number of estimations that show an increased excess variability. This number rises with the value of λ. For $\lambda = 10$ about the first and last three estimations are affected, for $\lambda = 500$ about the first and last eight estimations and for $\lambda = 10000$ more than the first and last 20 estimations show an increased *loss*. But while the number of affected estimations rises with increasing values of λ, the degree of the excess variability is worse for lower values of λ. For $\lambda = 10000$ the *loss* of the last estimation is about 0.15, while for $\lambda = 10$ it is about 0.6. However, the fact that the excess variability depends on λ is of subordinate importance here, as this paper only focuses on the standard case of $\lambda = 1600$.

3 Introducing a flexible penalization

Given the *loss function* it is possible to quantify and describe the excess variability. The next step is to find techniques to reduce this variability. To this point a flexible penalization for the HP-filter is introduced that allows selecting different values of λ for different points in time. According to Figure 3, the first and last ten or eleven estimations show an increased *loss*, when a HP-filter with $\lambda = 1600$ is applied. Consequently the penalization should rise for the estimations at the margins of the series. This can be done easily by changing the model framework of the HP-filter slightly. Given formula (3) the scalar λ has to be replaced by a vector $\boldsymbol{\lambda} \in \mathbb{R}^{T-2\times 1}$, where $\boldsymbol{\lambda}' = (\lambda_1, ..., \lambda_{T-2})$ and one constructs $\boldsymbol{K} = \text{diag}(\boldsymbol{\lambda})$. Then the HP-filter with a flexible penalization can be written as

$$\hat{\boldsymbol{\mu}} = (\mathbf{I} - \boldsymbol{\Delta}'\boldsymbol{K}\boldsymbol{\Delta})^{-1}\mathbf{y}. \tag{8}$$

As the HP-filter is equal to a penalized spline of order one with a truncated power basis and as many knots as observations (e.g. Proietti/Luati, 2007), it can be interpreted as a continuous connection of lines, where the slope of the lines changes at the points in time $t = 2, 3, ..., T-1$. To this regard the penalization regulates to what extend the slope can change at these points in time. High values of λ let the slope only change slightly which results in a smooth trend estimation whereas low values for the penalization allow large changes of the slope inducing a flexible trend function. λ_1 regulates the degree to what the slope of the trend function can change at $t = 2$ and λ_{T-2} determines the change of the slope at the point in time $T-1$. In general λ_t regulates the change of the slope at the point in time $t+1$. Thus it becomes obvious how the flexible penalization by the vector $\boldsymbol{\lambda}$ allows to vary the degree of smoothness over time.

A question that arises is how the penalization should increase to the margins. The general aim is to reduce the *loss* at the margins without increasing it in the middle of the data. To this point another criterion is defined, which will turn out to be suitable for this purpose. The penalization at the margins is increased such, that the cumulative *loss* of all estimations is minimized. This is reasonable as one is usually not just interested in the trend in the middle but in the whole series. Moreover it will be shown that this criterion will lead to a reduced excess variability at the margins without strongly affecting the estimations in the middle of the series. Defining the cumulative *loss* in dependence of $\boldsymbol{\lambda}$ as $L(\boldsymbol{\lambda})$, it can be written as:

$$L(\boldsymbol{\lambda}) = \sum_{t=1}^{T} l(t, \boldsymbol{\lambda}), \tag{9}$$

$$\text{where} \quad l(t, \boldsymbol{\lambda}) = \sum_{i=1}^{n} [g_m(\omega_i, \lambda) - g_t(\omega_i, \boldsymbol{\lambda})]^2 \cdot \delta. \tag{10}$$

Here $g_t(\omega_i, \boldsymbol{\lambda})$ is the gain of the t^{th} estimation for frequency ω_i using the flexible penalization and $g_m(\omega_i, \lambda)$ is the gain for the medium estimation when a single fixed penalization parameter is used. This criterion is subject to the conditions that the penalization is 1600 in the middle and that it rises to the margins of the series. Considering the *loss function* over all estimations suggests, that a linear increase of the penalization to the margins might be appropriate. As a consequence the last k values of $\boldsymbol{\lambda}$ rise by

$$\lambda_{T-2-k+j} = 1600 + \alpha j, \quad j = 1, ..., k. \tag{11}$$

The intercept can be seen as given, as λ is set to 1600 for the estimations in the middle. As the penalization needs to increase to both margins the first k values of the penalization are defined as

$$\lambda_1 = \lambda_{T-2}, \lambda_2 = \lambda_{T-3}, ..., \lambda_k = \lambda_{T-1-k}. \tag{12}$$

Besides a linear increase of the penalization also other functions like a quadratic or cubic increase or a polynomial of degree two could be assumed. However the results would hardly change, but especially the polynomial function would make the calculation much more expensive. Minimizing (9) with respect to (11) and (12) can be done by algorithms like fisher scoring or Newton-Raphson. However the minimization by these algorithms can only be done for α as k is an integer. Thus one has to minimize the cumulative *loss* for different k's and use the one that yields the lowest value of $L(\boldsymbol{\lambda})$.

Applying this algorithm to a simulated time series with 100 observations yields $k = 27$ and $\alpha = 1294.72$. The penalization parameter for the last 27 λ's then rises by $\lambda_{71+j} = 1600 + 1294.72 \cdot j, \quad j = 1, ..., 27$, while $\lambda_1, ..., \lambda_{27}$ are defined according to (12) (This paper focuses on the case of $\lambda = 1600$ but a table with corresponding values for k and α for other values of λ is provided in the appendix). Table 1 shows the *loss* for the fixed and the flexible penalization for the middle of the data and the margin ($50^{th}$and 100^{th} estimation) as well as $L(\boldsymbol{\lambda})$.

Table 1: *loss* for series with 100 observations

penalization	**50^{th}estimation**	**100^{th}estimation**	**$L(\boldsymbol{\lambda})$**
fixed	0	0.23956	1.76382
flexible	0.00015	0.09078	1.16872

As Table 1 shows, the flexible penalization reduced the *loss* for the 100^{th} estimation as well as $L(\boldsymbol{\lambda})$, while it increased the one for the 50^{th} estimation only slightly. $L(\boldsymbol{\lambda})$ could be reduced by around 34 percent and for the 100^{th} estimation the *loss* even declined by 62 percent. This indicates, that the excess variability can be reduced without affecting the estimations in the middle of the data. However, to get a complete picture of how the flexible penalization changes the gain function of all estimations, Figure 5 displays the *loss* for the whole series for the fixed and the flexible penalization.

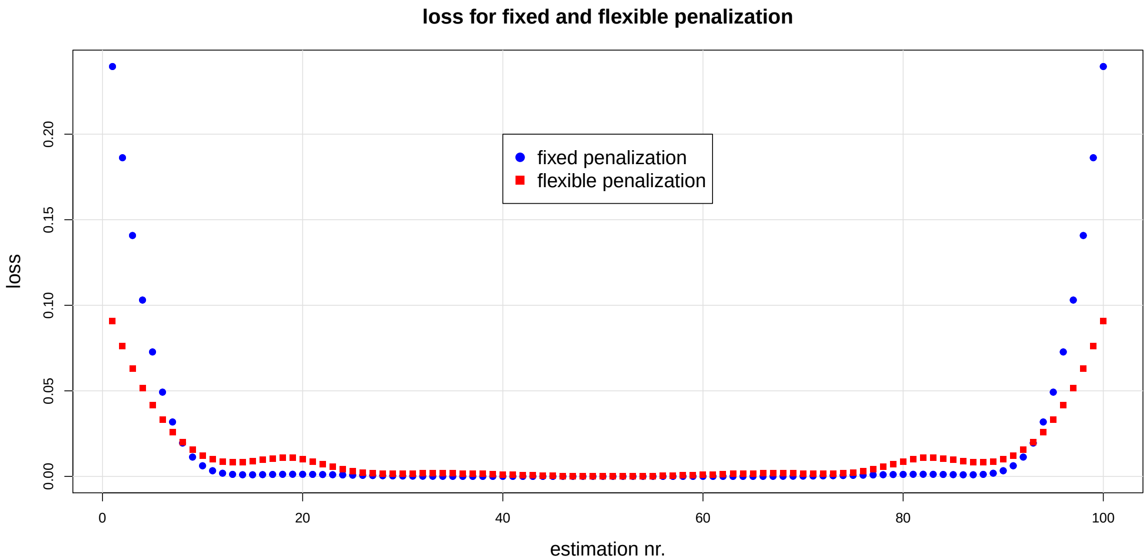


Figure 5: *loss function* for a fixed and a flexible penalization

Clearly the *loss* strongly decreased for about the first and last six estimations. While the *loss* was strongly reduced at the margins, it was slightly increased by the flexible penalization for about the estimations 10-20 and 80-90. However the decrease at the margins outweighs this slight increase which results in the strong decrease of the cumulative *loss*. Thus the flexible penalization combined with spectral analysis offers a tool to reduce the excess variability without strongly affecting the trend in the middle of the data (note that a simpler approach to reduce the excess variability of the HP-filter was already suggested by Bruchez (2003)). The next section will apply this method to some real time series and shortly discuss the implications that arise when one switches from a fixed to a flexible penalization.

4 Empirical application

In this section the flexible penalization is applied to empirical time series in order to point out how results might change when one allows for a flexible instead of a fixed penalization. First of all the seasonally adjusted quarterly real GDP of Switzerland is considered[1]. The data start in the first quarter 1980 and end in the third quarter 2013 so that there are 135 observations. The minimization of the cumulative *loss* $L(\boldsymbol{\lambda})$ with respect to α and k yielded $\alpha = 1304.22$ and $k = 27$. The trend is also estimated by a HP-filter with a fixed penalization of $\lambda = 1600$. The results are shown in Figure 6:

[1]The data are from the Swiss State Secretariat of Economic Affairs, http://www.seco.admin.ch.

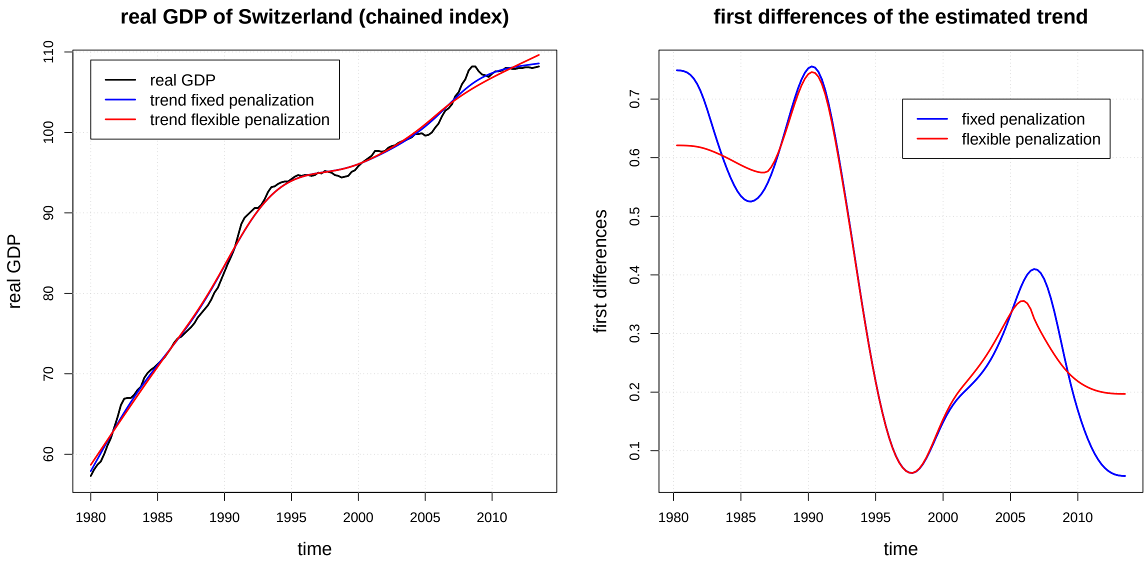


Figure 6: Estimated trend for the Swiss real GDP and first differences of the trend

As one can see the growth of the trend according to the fixed penalization has strongly decreased since about 2007. The trend of the flexible penalization exhibits only a slight decline of its growth since 2007 and its growth rate is clearly above the case of the fixed penalization for the last three years. This can also be seen when the first differences on the right plot are considered. Furthermore the trend of the flexible penalization is clearly above the one of the fixed penalization for the most recent years. A further interesting feature of the flexible penalization becomes obvious, when the resulting estimates of the business cycle are considered.

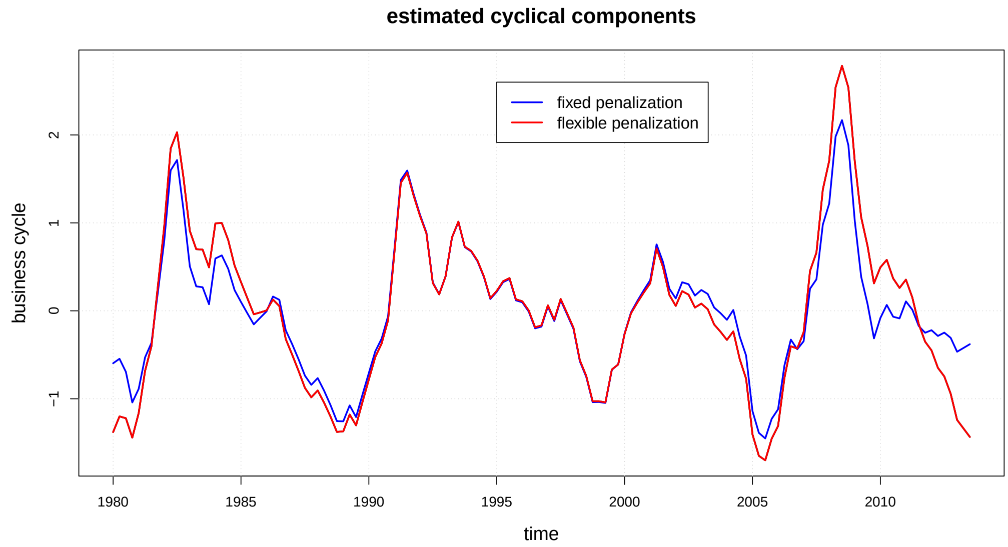
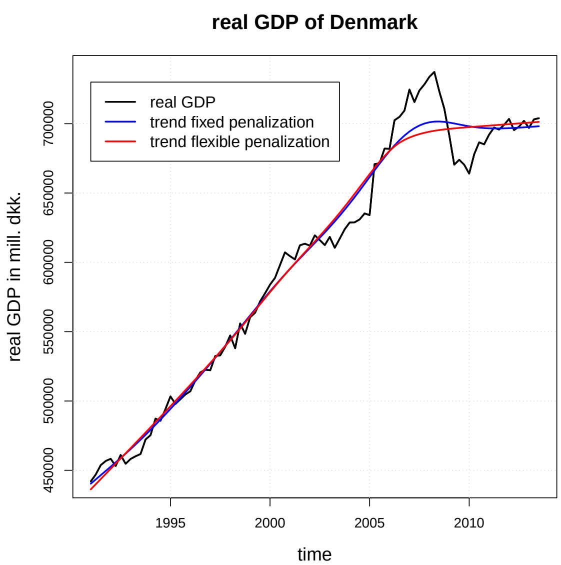


Figure 7: Estimated business cycles for fixed and flexible penalization

As Figure 7 shows, the output gap at the end of the series is much smaller for the fixed penalization. This smaller output gap is a direct consequence of the excess variability. As Figure 1 shows, the filter weights for estimations at the margins are very high for the last observation. Thus the last observation is the most influential for these estimations which induces that the estimations at the margin are deterred to the value of the last observation. As a result in this example the excess variability in trend leads to a rather too low output gap at the margin of the series. The cyclical component of the flexible penalization instead shows a far larger output gap at the end of the series, as the excess variability could be reduced.

Next the seasonally adjusted quarterly real GDP of Denmark is considered[2]. The data start in the first quarter 1991 and end in the third quarter 2013 so that there are 91 observations in this series. Minimizing the cumulative *loss* $L(\boldsymbol{\lambda})$ yields $\alpha = 1242.48$ and $k = 27$. The trend of the real GDP is estimated using both the flexible and a fixed penalization. The results are shown in the left plot of Figure 8.



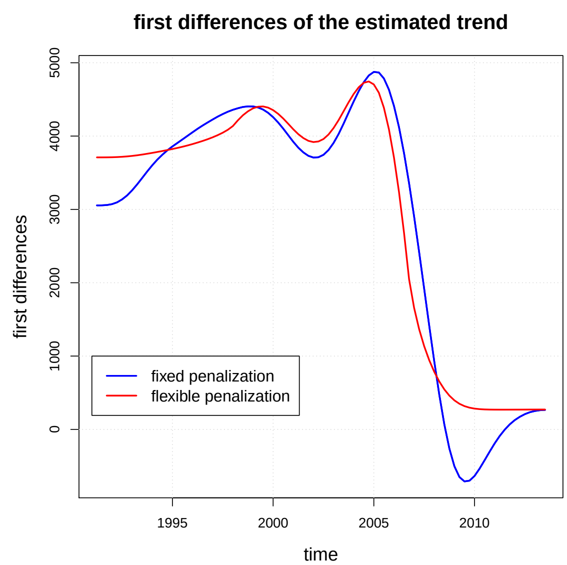


Figure 8: Estimated trend for the Danish real GDP and first differences of the trend

There are clear differences between the estimations. In both cases the growth rate of the trend decreases from about 2005, but the decline is much larger for the fixed penalization. The right plot of Figure 8 shows the first differences of the trend estimations. Here the differences become even more obvious. One can see that according to the fixed penalization the growth of the trend even tuned negative for the years 2008-2012. In contrast the trend growth rate according to the flexible penalization declined from the year 2005, but seems to have stabilized on a lower level during the last three years.

[2]The data are from the national census bureau of Denmark, http://www.statbank.dk.

5 Conclusion

The Hodrick-Prescott filter is the probably most widespread instrument for trend estimation in economics. Compared to the competing Baxter-King filter (Baxter/King, 1999) it offers the advantage of yielding estimates for the most recent periods. However, the fact that the filter weights strongly change at the margins leads to an increased excess variability for these periods. This means, that compared to the estimation in the middle of the data the trend at the margins is too volatile. Especially as researchers are predominantly interested in the most recent periods, the excess variability turns out to be a serious problem of the Hodrick-Prescott filter. An existing method to overcome this problem is to use ARIMA models in order to prolong the time series by adding forecasts at the margins. Nevertheless, the predictions exhibit failures that increase with a rising forecast horizon. As for $\lambda = 1600$ more that ten periods have to be predicted, this method seems to be of limited practicability.

This paper combined spectral analysis with a flexible penalization in order to reduce the excess variability at the margins. To this point the $loss$, i.e. the squared deviation of the gain function from the one in the middle of the series was used as a measure to describe and quantify the excess variability. To reduce the increased volatility of the estimations at the margins the penalization was allowed to increase linearly to the margins. The exact rise of the penalization was determined such that the cumulative $loss$ is minimized. In this regard it was shown that this criterion not only leads to a lower cumulative $loss$, but that it reduces the excess variability at the margins without strongly affecting the estimations closer to the middle of the series.

To show the empirical implications that can arise when the flexible penalization is used instead of a fixed one the HP-filter with a flexible penalization was used to estimate the trend of the real GDP of Switzerland and Denmark. It was shown that in both cases the estimated trend according to the flexible penalization considerably differed from the trend estimation with the fixed penalization. Especially the estimation for Switzerland using the flexible penalization showed, that the current data of the Swiss real GDP do not allow to conclude that the trend growth rate has decreased immensely since 2007, like the HP-filter with the fixed penalization suggests.

Thus, this paper offered an approach to improve estimations at the margins. Although the excess variability could not be completely eliminated it was reduced by more than 62 percent for the last estimation. As the empirical examples showed, the results of the flexible penalization can strongly differ from the standard approach with a fixed λ of 1600. Given this, the flexible penalization might be an interesting tool for researchers as it allows to increase to precision of the estimations for the most recent periods.

A Flexible penalization for different values of λ

The following table provides the corresponding values of k and α for different values of λ. Please note that the values can vary slightly for series of different length. The table also shows the length of the series to which the values refer, which is denoted as T_{ref}. However, in most cases the results only change slightly when these values are used for series with a length that deviates from the reference length in this table as long as the series are not too short.

Table 2: values for k and α for different values of λ

λ	k	α	T_{ref}	λ	k	α	T_{ref}
5	5	19	100	1000	24	880	125
10	6	44	100	1250	25	1165	150
20	8	47	100	1500	26	1414	150
30	9	66	100	1750	27	1617	150
40	10	72	100	2000	27	2279	150
50	10	126	100	2500	28	3198	150
60	10	210	100	3000	29	4073	150
70	12	99	100	3500	31	3713	150
80	12	134	100	4000	32	4207	150
90	12	179	100	4500	33	4606	150
100	13	144	100	5000	34	4904	150
200	15	325	100	6000	38	3617	200
300	17	403	100	8000	41	4452	200
400	19	412	100	10000	43	5611	200
500	20	513	100	12000	46	5673	200
600	21	590	125	14000	47	7133	250
700	21	853	125	16000	50	6660	250
800	22	881	125	18000	51	7752	250
900	23	887	125	20000	52	8796	250

B References

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

Bruchez, P. A. (2003): *"A Modification of the HP-filter - Aiming at Reducing the End-Point Bias,"* Swiss Federal Finance Administration, Working Paper.

Colgley, T., Nason, J. (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, 253-278.

Crainiceanu, C., Ruppert, D., Carroll, R. (2005): *"Spatially adaptive Bayesian P-splines with heteroscedastic errors,"* Working Paper.

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

Flaig, G. (2012): *"Why we should use high values for the smoothing parameter of the Hodrick-Prescott filter,"* CESifo Working Paper Series No. 3816.

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

Harvey, A. C. (1993): *"Time Series Models,"* Harvester Wheatsheaf, Hertfordshire.

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

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

King, R., Rebelo, S. (1993): *"Low Frequency filtering and real business cycles,"* Journal of Economic Dynamics and Control 17, 207-231.

Maravall, A., del Rio, A. (2001): *"Time Aggregation and the Hodrick-Prescott-Filter,"* Banco de Espana Servicio de Estudios, Documento de Trabajo n. 0108.

McCallum, B. T. (2000): *"Alternative monetary policy rules: A comparison with historical settings for the United States, the United Kingdom and Japan,"* Economic Quarterly of the Federal Reserve Bank of Richmond, 86, 49-79.

McElroy, T. (2008): *"Matrix Formulas for Nonstationary ARIMA Signal Extraction,"* Econometric Theory 24, 988-1009.

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.

Proietti, T. (2007): *"Signal Extraction and Filtering by Linear Semiparametric Methods,"* Computational Statistics and Data Analysis 52, 935-958.

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

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.

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.

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

Tödter, K. H. (2002): *"Exponential Smoothing as an Alternative to the Hodrick-Prescott Filter,"* Contributions to Modern Econometrics-From Data Analysis to Economic Policy. Boston: Kluwer Academic Publishers, S. 223-237.