

Research Report

Department of Statistics

No. 2007:1

A heteroscedasticity removing filter

**Pär Stockhammar
Lars-Erik Öller**

A Simple Heteroscedasticity Removing Filter

Pär Stockhammar and Lars-Erik Öller

Department of Statistics, Stockholm University
S-106 91 Stockholm, Sweden
E-mail: par.stockhammar@stat.su.se

Abstract

In this paper variance stabilizing filters are discussed. A new filter with nice properties is proposed which makes use of moving averages and moving standard deviations, the latter smoothed with the Hodrick-Prescott filter. This filter is compared to a GARCH-type filter. An ARIMA model is estimated for the filtered GDP-series, and the parameter estimates are used in forecasting the original series. These forecasts compare well with those of ARIMA, ARFIMA and GARCH models based on the unfiltered data. The filter does not colour white noise.

Keywords: Economic growth, heteroscedasticity, variance stabilizing filters, the Hodrick-Prescott filter.

1. Introduction

Data transformations are made in order to facilitate analysis of empirical time series. There are a number of reasons why one might want to remove heteroscedasticity before modelling. For one thing, it saves on parameters. Another reason is the fact that most time series models require stationarity. Constant mean and variance are necessary requirements for (weak) stationarity. In case the variance is proportional to the level of the series, a logarithmic transformation may make the series both homoscedastic and stationary in variance. But many time series do not have constant, or even stationary variance even after transformations. Often heteroscedasticity is simply ignored and there does not seem to exist heteroscedasticity removing filters in the literature. Handling this problem is still a hot topic, see e.g. Baltagi et. al (2009). Giordani and Villani (2010) suggested the locally adaptive signal extraction and regression (LASER) model to capture the dynamics of the signal and noise. The mean and variance of these components are allowed to slowly or abruptly shift and the model (using a normal mixture distribution for each component) is then able to capture these

movements.

In this paper we present a new filter designed to remove heteroscedasticity. This is achieved by dividing the time series with a moving average of its standard deviations (STDs), smoothed by a Hodrick-Prescott filter (HP). Here we apply the filter to the logarithmic, quarterly and seasonally adjusted US GDP-series. The same series of the UK and Australia are also analyzed. All the above time series are found to be significantly heteroscedastic. After filtering no heteroscedasticity remains. Moreover, white noise is not coloured.

Within the model building framework, there exists numerous ways to handle heteroscedasticity such as weighted least squares (WLS) or generalized least squares (GLS) in regression models, autoregressive conditional heteroscedasticity (ARCH) and generalized ARCH (GARCH) models for time series data. The most straightforward way to remove heteroscedasticity in the GDP series above is to divide the heteroscedastic series by the conditional volatility estimated from ARCH/GARCH models or from any of their many generalizations. A comparison and a discussion of the two approaches will be pursued in this paper.

Despite the heteroscedasticity in the first differenced logarithmic (Diff ln) series, ADF tests do not signal any unit root. This can lead to a statistically correct, but inefficient use of e.g. ARIMA or ARFIMA models. Hess and Iwata (1997) conclude that there is no model that better replicates business cycle features than the simple ARIMA(1,1,0) model. This was contested by Candelon and Gil-Alana (2004) who considered fractionally integrated models, and showed that ARFIMA models even more accurately describe the business cycle characteristics in the US and the UK. After applying the proposed filter, an adequate ARIMA-model is estimated for the filtered GDP-series, and the parameter estimates are then used in forecasting the original time series. The forecasts are compared with those from ARIMA, ARFIMA and GARCH models estimated from unfiltered data, showing that ARFIMA is not the best model. The effect of the inefficiency is measured.

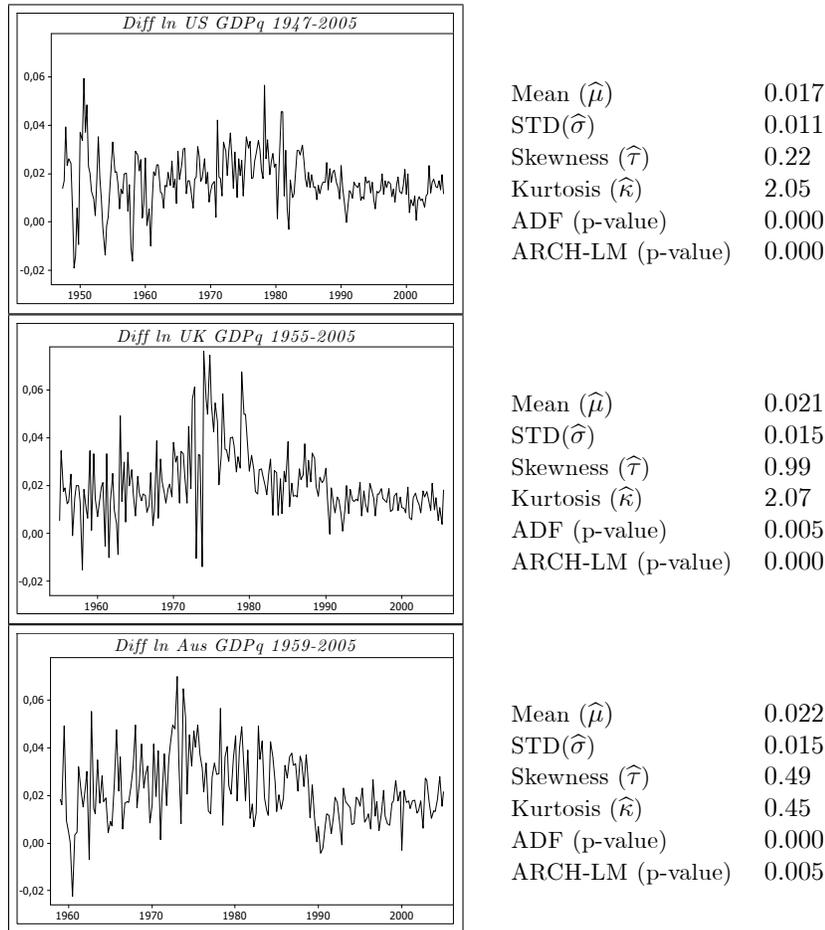
The data are presented in Section 2 and the effects of heteroscedasticity are discussed in Section 3. Section 4 contains a discussion about the filters, where we also test for possible side effects. Section 5 contains the forecast comparison and Section 6 concludes.

2. The data

The seasonally adjusted US GDP $_q$ (quarterly) series 1947-2005 ($n = 236$ observations) can be found on the website of *Bureau of Economic Analysis*, www.bea.gov. The seasonally adjusted UK GDP $_q$ 1955-2005 ($n = 204$), and the Australian GDP $_q$ series 1959-2005 ($n = 188$) were copied from the websites National Statistics (www.statistics.gov.uk) and of OECD (www.oecd.org), respectively. The

Diff ln GDP quarterly time series together with its first four estimated moments are shown in Figure 2.1. The results of the augmented Dickey-Fuller (ADF) test for a unit root and the ARCH-Lagrange multiplier (ARCH-LM) test for homoscedasticity are also included.

Figure 2.1: The Diff ln GDP series and their properties



The skewness is significantly nonzero in the UK and Aus series and significant leptokurtocity appears in the US and UK series. Unit root tests are highly sensitive to heteroscedasticity. Hamori and Tokihisa (1997) showed that a permanent STD shift strongly affects the size of Dickey-Fuller type tests. The effect of a single break in STD on the ADF test has been analyzed by Kim et al. (2002), who reported risks of over-rejection of the null hypothesis in the presence of a negative break. This builds on the study by Kim and Schmidt (1993) who

showed that Dickey-Fuller tests tend to reject too often in the presence of conditional heteroskedasticity. The discussion was extended by Cavaliere (2004), who came to the same conclusion for other commonly used unit root tests. In our study, the null hypothesis of a unit root could not be rejected in all original series, but rejected for the Diff ln series, signalling stationarity. However the persistence in the series in Figure 2.1 may hint at an integrating order between zero and one, which may be hard to detect due to heteroscedasticity.

3. Effects of heteroscedasticity

Often nonstationary economic time series can be made stationary simply by differencing. But this can remedy only nonstationarity in mean - nonstationarity in variance must be handled in other ways.

Let the variance of a nonstationary process change with the level, $Var(y_t) = cf(\mu_t)$, for any positive proportionality constant c , and where f is an increasing function of the time varying level μ_t . Then it is possible to find a transformation T so that $T(y_t)$ has constant variance by approximating the function by a first order Taylor series around μ_t

$$T(y_t) \approx T(\mu_t) + T'(\mu_t)(y_t - \mu_t).$$

Now

$$Var [T(y_t)] \approx [T'(\mu_t)]^2 Var(y_t) = cf(\mu_t) [T'(\mu_t)]^2.$$

For the variance of $T(y_t)$ to be constant the transformation must be chosen so that

$$T'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}},$$

that is

$$T(\mu_t) = \int \frac{1}{\sqrt{f(\mu_t)}} d\mu_t.$$

If, for example, $Var(y_t) = c^2\mu_t^2$, then $T(\mu_t) = \int \frac{1}{\sqrt{\mu_t^2}} d\mu_t = \ln \mu_t$. Hence, a logarithmic transformation of the series will have a constant variance.

If instead the variance of the series is linearly proportional to the level so that $Var(y_t) = c\mu_t$, then the square root transformation $\sqrt{y_t}$ will produce a constant variance. More generally, the Box-Cox transformation (Box and Cox, 1964) that includes the logarithmic and the square root transformations as special cases, is often used to stabilize the variance. However the transformations mentioned above are only defined for positive series, and more importantly, what to do when the standard deviation is not a function of the series?

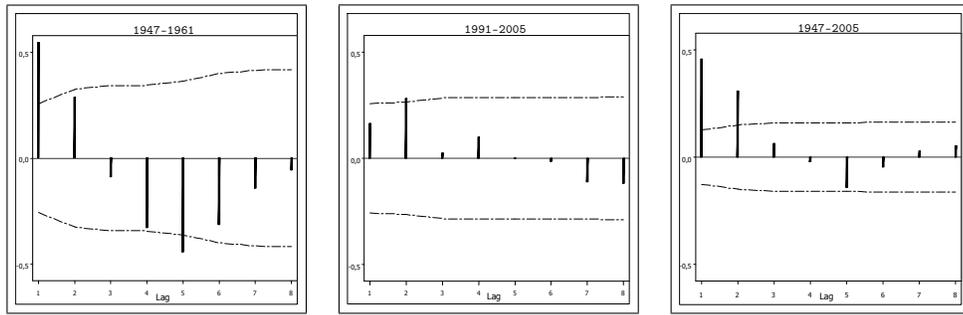
In regression it is well known that OLS-estimates are not efficient in the presence

of heteroscedastic disturbances. The problem can be solved using the standard theory of GLS by simply dividing each observation and each explaining variable by an estimate of the varying standard deviation of the disturbances. OLS estimates of the transformed model would then be efficient.

In the univariate case one likes to preserve the dynamic structure (autocorrelation) while making the series homoscedastic. But, here too, the non-constant standard deviation σ_t is unknown and must be estimated. With just one realisation of the series this can not be done. A way out is to estimate σ_t recursively using a window of observations¹. As in the GLS case, dividing by appropriate observation weights would produce estimates that are BLUE.

Consider Figure 3.1 showing the autocorrelation function (ACF) of three time periods of the US GDP-series, the first and last 15 years, and the entire series.

Figure 3.1 ACF of Diff ln US GDP_q



There are obvious similarities in autocorrelation structure between the first and most volatile period (1947-1961), and the entire series (1947-2005). The most recent period (1991-2005), which is crucial for forecasts, does not seem to have much influence on the ACF estimates for the entire period. There has been a change in the autocorrelation structure, but with equal observation weights the model estimates will mainly be based on the old structure, because of the high volatility in the beginning of the time series.

4. Heteroscedasticity filters

Prior to heteroscedasticity filtering, the series is stationarised and normalised to symmetrically vary around zero. Otherwise the filtered GDP series would contain a trend. The Diff ln series in Figure 2.1 all have a nonzero mean and despite the rejection of a unit root, they exhibit slowly changing levels. This

¹Define the variance of a moving window of length $2\eta + 1$, with η even:

$$Var(y_t) = \frac{1}{2\eta+1} \sum_{i=-\eta}^{\eta} (y_t - \mu_t)^2, \text{ where } \mu_t = \frac{1}{2\eta+1} \sum_{i=-\eta}^{\eta} y_t$$

hints at the possibility of a double root, or the true integrating order could lie somewhere between $I(1)$ and $I(2)$, suggesting fractionally integrated models, such as ARFIMA. This is also supported by Candelon and Gil-Alana (2004) who concluded that the US and UK GDP series are integrated of order around $I(1.4)$. In the nine cases in the subsequent Section 5, the root is (along with the other parameters) estimated to lie between 1.27 and 1.47. Taking just one difference is not enough to extract all trend whereas taking another difference over-differences the data. One solution to the problem is simply to take a fractional difference of degree $d = 1.4$ prior to heteroscedasticity filtering. The local trends in the series will then be close to eliminated. In general, yet another parameter (d) needs to be estimated and results based on fractionally integrated time series are difficult to interpret.

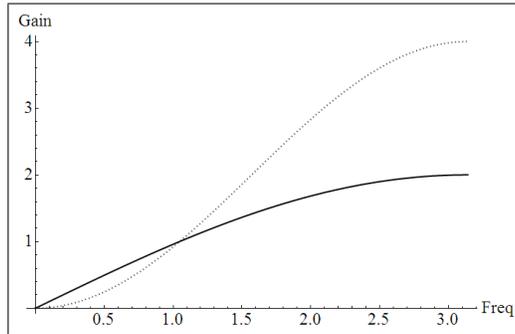
4.1 A simple filter

Transforming the series by subtracting from each (Diff ln) observation a local mean is a tempting alternative. It is possible to capture integration orders between $I(1)$ and $I(2)$ using one of the following two operations:

$$\begin{aligned}
 (a) \quad z_t^{(a)} &= \Delta y_t - \sum_{\tau=t-\eta}^{t+\eta} \Delta y_\tau / k, & t = \eta + 1, \eta + 2, \dots, n - \eta \\
 (b) \quad z_t^{(b)} &= \Delta y_t - \sum_{\tau=t-\eta}^{t+\eta} \Delta y_{\tau-1} / k, & t = \eta + 2, \eta + 3, \dots, n - \eta + 1
 \end{aligned}
 \tag{4.1}$$

where the centered $\Delta y_t = y_t - y_{t-1}$, y_t is the log GDP series at time t , k (odd) is the window length, $\eta = (k - 1)/2$ and even. Using different values on η in (4.1), different degrees of integration are achieved. There are two extremes. For $\eta = (n - 1)/2$, the term $\sum_{\tau=t-\eta}^{t+\eta} \Delta y_\tau / k$ equals $\overline{\Delta y}$. The other extreme appears when k equals one, that is $\eta = 0$. Operation (4.1b) is used only in the latter case and is equivalent to a the second difference operation, $\Delta^2 y_t$. The choice of η depends on the series studied. If it is close to $I(1)$ then you should just choose η close to $(n - 1)/2$, and if the series is close to $I(2)$ then choose $\eta = 0$ in (b) or a small value on η in (a). No matter what choice of η , (4.1) is a high-pass filter in that it removes the low frequency movements of the series. The gain is the change in the output when a step change of unit size hits the input. Figure 4.1 shows the gain for the two extremes.

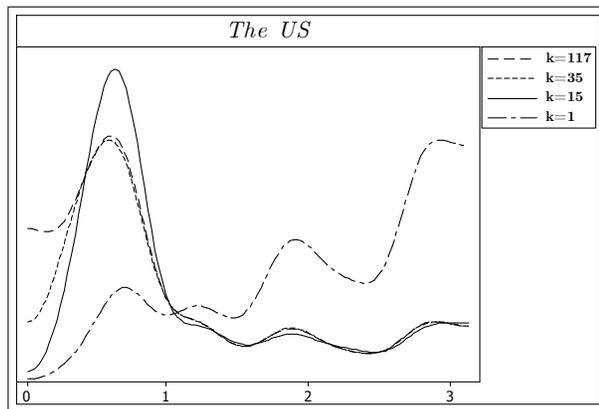
Figure 4.1: The gain functions of the special cases one difference, $\eta = (n - 1)/2$ (solid line) and two differences, $\eta = 0$ (dotted line)

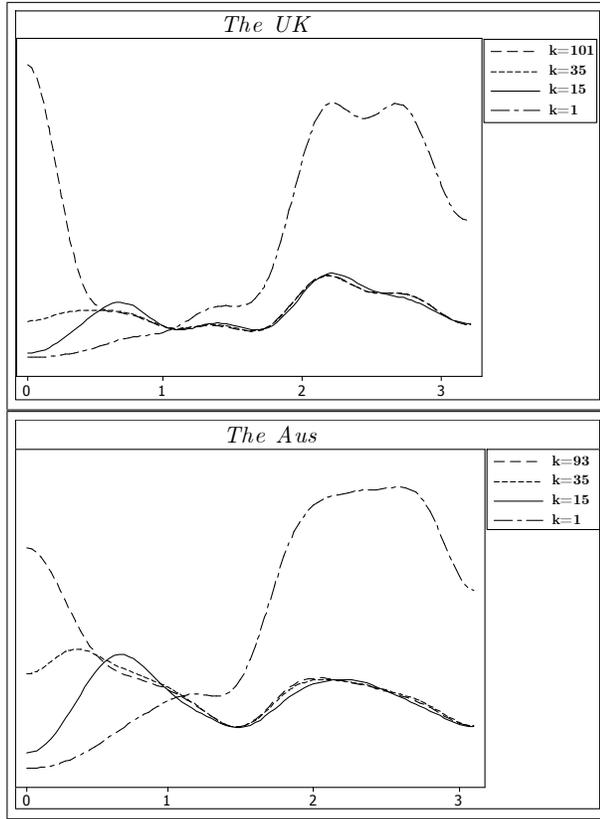


This transformation can be generalized by raising z_t to the power d , $(z_t^{(i)})^d$ where $i = a, b$. This enables us to handle integrating orders below one and above two. However, this is not needed for the series studied in this paper.

Figure 4.2 shows the estimated spectral densities of the (4.1) for the Diff ln US, UK and Aus data, using window length $k \approx n/2$, $k = 35$, $k = 15$ and $k = 1$. In all cases a Parzen window was used to smooth the spectral densities.

Figure 4.2: The spectrum of the ln US, UK and Aus data





For the extreme of just one difference one gets a spectrum dominated by low frequency variations, due to the persistence in these series. On the contrary, using (4.1b) with $k = 1$ ($\Delta^2 y_t$) removes all variation at the zero frequency in these series. Neither of the two extremes is very attractive to use here, the dominant low or high frequency properties overshadow frequencies in between. The remaining alternatives all have very similar high frequency properties. As expected when the windows get shorter the low frequencies are transferred to higher frequencies. The high frequency variations resulting from even shorter windows is too dominating. It seems that window length $k = 15$ is a good middle course which succeeds in detrending the data without removing business cycle features. Using $k = 15$, all three filtered spectral densities are bimodal with peaks at around three and ten quarters. The first four moments of $z_t^{(a)}$ from (4.1), using $k = 15$, are reported in Table 4.1.

Table 4.1: The moments of $z_{t,US}^{(a)}$, $z_{t,UK}^{(a)}$ and $z_{t,Aus}^{(a)}$ using $k = 15$

	US	UK	Aus
Mean ($\widehat{\mu}$)	0.000	0.000	0.000
STD($\widehat{\sigma}$)	0.011	0.011	0.012
Skewness ($\widehat{\tau}$)	0.034	-0.459	0.123
Kurtosis ($\widehat{\kappa}$)	2.362	4.681	0.932

The skewness is significantly nonzero in the UK, and significant leptokurtocity appears in all series, cf. Figure 2.1. Moreover the series are still significantly heteroscedastic but stationary.

Removing the heteroscedasticity is a matter of dividing (4.1) by estimates of the changing volatility. It does not seem reasonable to assume that the volatility changes abruptly all the time - smoothing is necessary as a compromise between contrafactual constant STD and unrealistically large volatility changes from one quarter to the next. Hence we assume that the variance is slowly evolving over time. Thus, a logical estimate of the volatility at time t is

$$HP^{(\gamma)} \left\{ \sqrt{\frac{\sum_{\tau=t-\nu}^{t+\nu} (z_{\tau}^{(i)})^{2d}}{2\nu}} \right\}, \quad i = a, b \quad (4.2)$$

where $\nu = (l - 1)/2$ and l is the window length which might not be equal to the window length, k , in (4.1). In this study however, $k = l$. $HP^{(\gamma)}$ is the Hodrick-Prescott (1997) filter designed to decompose a macroeconomic time series into a nonstationary trend component and a stationary cyclical residual.

Given a time series x_t (in this case the varying variance), the decomposition into unobserved components is

$$x_t = g_t + c_t,$$

where g_t denotes the unobserved trend component at time t , and c_t the unobserved cyclical residual at time t . Estimates of the trend and cyclical components are obtained as the solution to the following minimization problem

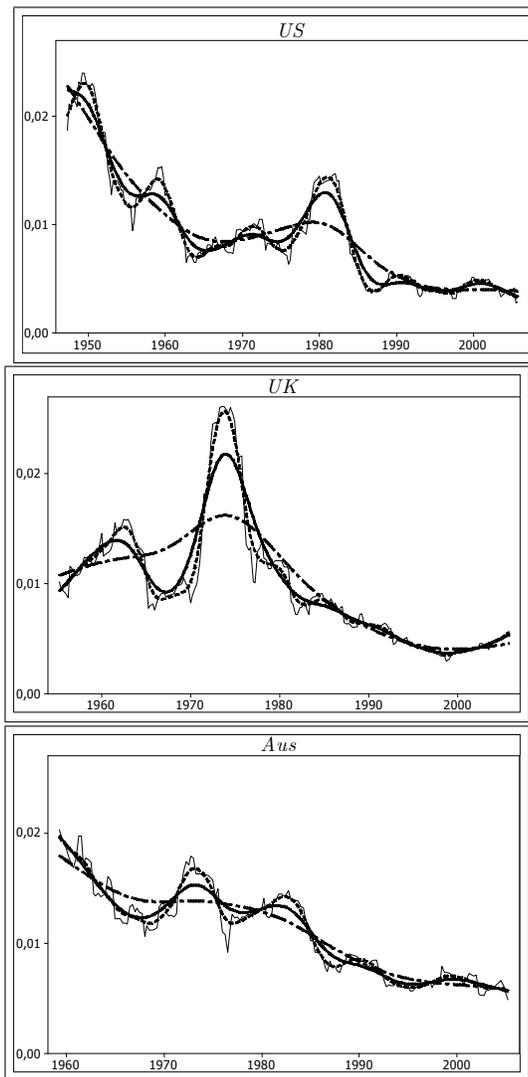
$$\min_{\{g_t\}_{t=1}^n} \left\{ \sum_{t=1}^n c_t^2 + \gamma \sum_{t=3}^n (\Delta^2 g_t)^2 \right\}, \quad (4.3)$$

where $\Delta g_t = g_t - g_{t-1}$ and g_{\min} is the HP-filter. The first sum of (4.3) expresses the closeness between the HP trend and the original series, while the second sum represents the smoothness of the trend. The positive smoothing parameter γ controls the weight between the two components and is thereby a measure of the

signal-to-noise variance ratio. As γ increases, the HP trend becomes smoother and vice versa. Note that the second sum, $(\Delta^2 g_t)$, is an approximation to the second derivative of g at time t .

Figure 4.3 shows the moving STDs and illustrates the effect of the HP-filter for different values of γ .

Figure 4.3: The moving STDs (thin solid line) and the HP trend using $\gamma = 100$ (dashed), $\gamma = 1\ 600$ (thick solid) and $\gamma = 50\ 000$ (dashed/dotted)



If we would not have made the mean correction in (4.1) a positive trend would

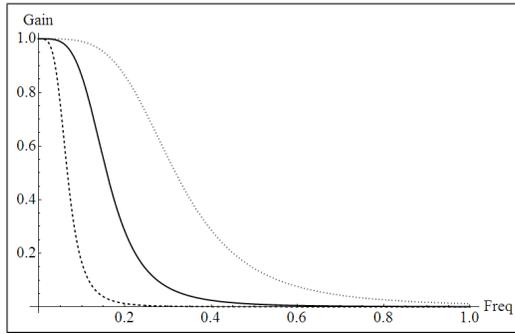
have resulted in \tilde{z}_t below, cf. Figure 2.1 and 4.3. For quarterly data a commonly used value is $\gamma = 1\ 600$, originally proposed in Hodrick and Prescott (1997)². In this study, $\gamma = 1\ 600$ accords well with the principle "not too rugged - not too smooth".

Dividing (4.1) by (4.2) and multiplying by the overall STD of the entire period, s_y , and adding the corresponding arithmetic average \bar{y} , we get the proposed heteroscedasticity removing filter

$$\tilde{z}_t = s_y \left[\frac{(z_t^{(i)})^d}{HP^{(\gamma)} \left(\sqrt{\sum_{\tau=t-\nu}^{t+\nu} (z_\tau^{(i)})^{2d} / 2\nu} \right)} \right] + \bar{y}, \quad (4.4)$$

where $t = \max[k - \eta, l - \nu], \max[k - \eta + 1, l - \nu + 1], \dots$, \tilde{z}_t is the filtered series and $i = a, b$. Dividing (4.1) by the moving STDs directly remains an alternative hypothesis to be considered later on in this section. Whatever the choice of γ in (4.4), the trend component, g_t , of the HP filter is a low-pass filter, see Figure 4.4.

Figure 4.4: The gain functions of the HP filter using $\gamma = 100$ (dotted line), $\gamma = 1600$ (solid line) and $\gamma = 50\ 000$ (dashed line)³



This means that for any parameter values, (4.4) is a high-pass filter. The filter is ad hoc in the same sense as the X11 seasonal adjustment. It is well known that seasonal adjustment filters may introduce spurious autocorrelation (as does X11), see e.g. Wallis (1974). This would be a serious disadvantage of the filter and can be tested by feeding white noise into (4.4). The above filter was applied on 10 000 simulated i.i.d $N(0, 1)$ series with 200 observations each, the

² "...a 5 percent cyclical component is moderately large, as is a one-eighth of 1 percent change in the growth rate in a quarter", yielding $\sqrt{\lambda} = 5/(1/8)$, $\lambda = 1600$.

³The gain function of the HP filter is $G(w, \lambda) = \frac{1}{1+4\lambda(1-\cos w)^2}$, where w is the frequency.

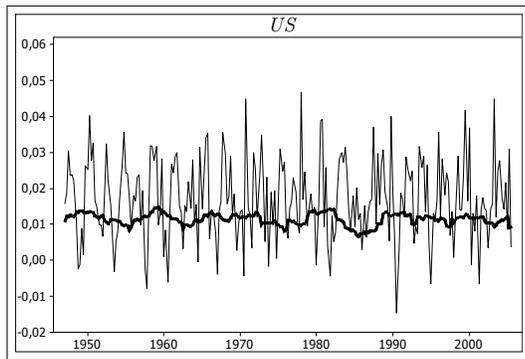
approximate length of the GDP-series in this study. The results are shown in Table 4.2.

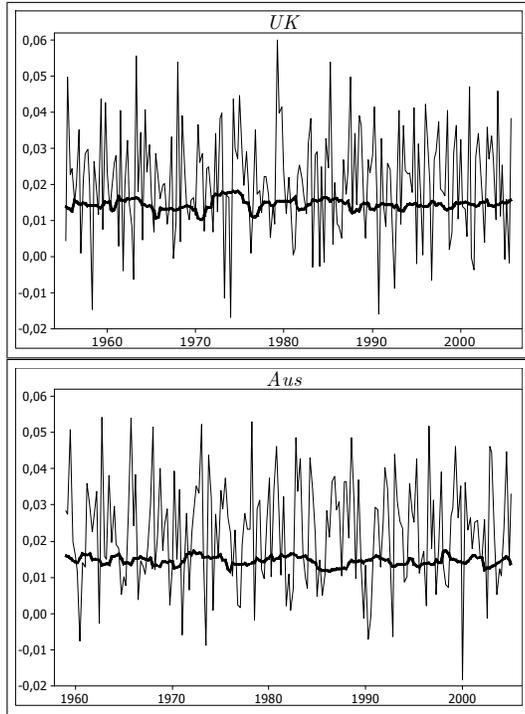
Table 4.2: Filter effects on white noise

	White noise	$\gamma = 1\ 600$	$\gamma = 10\ 000$	$\gamma = 100\ 000$
Mean	0.00005	0.00001	0.00002	0.00001
STD	0.99968	0.99244	0.99468	0.99725
Skewness (τ)	0.00002	0.00046	0.00035	0.00021
Kurtosis (κ)	2.99983	2.94535	2.96130	2.98420
Jarque-Bera	491/10000	488/10000	490/10000	504/10000
$Q_{0.05}^*(12)$	486/10000	651/10000	589/10000	507/10000
$Q_{0.05}^*(24)$	489/10000	691/10000	612/10000	521/10000

The first four moments in Table 4.2 are close to identical for the white noise series, unfiltered and filtered. The size of the Jarque-Bera test for normality is also correct at confidence level 0.05, the size of the Ljung-Box autocorrelation tests $Q_{0.05}^*(12)$ and $Q_{0.05}^*(24)$ is slightly too large, but no major distortion appears. Still, for $\gamma = 1\ 600$, the right hand tail of the χ^2 -distribution is slightly too thin. Thus, for $\gamma = 1\ 600$ a threshold value of 0.065 ($\chi_{0.065}^2(12) = 19.85$) is more appropriate for the 0.05 significance level, for 12 lags, and 0.07 ($\chi_{0.07}^2(24) = 35.24$) for 24 lags. When $\gamma \rightarrow \infty$, the HP-filter becomes linear, exactly preserving the distribution. Figure 4.5 shows the filtered series \tilde{z}_t with $k = l = 15$, $d = 1$ and $\gamma = 1\ 600$. Series of moving STDs (thick solid lines) are also included (using window length 15), in order to elucidate the effect of the filtering.

Figure 4.5: The filtered series





Filtered series using (4.4) but without HP smoothing of the STDs

$$\tilde{z}_t^{(\text{no HP})} = s_y \left[\frac{\left(z_t^{(i)} \right)^d}{\sqrt{\sum_{\tau=t-\nu}^{t+\nu} \left(z_\tau^{(i)} \right)^{2d} / 2\nu}} \right] + \bar{y}$$

are very similar. For the sake of completeness, the two alternatives will be compared. Figure 4.5 shows that both the mean and the STD look stable in the filtered series. As expected these series pass stationarity and homoscedasticity tests, see Table 4.3. Because of the low power of the ADF and ARCH-LM test in possible concurrent effects of levelshifts and heteroscedasticity, these tests are supplemented by the results of the Phillip-Perron (P-P) test for a unit root and the Breusch-Pagan(B-P) test of homoscedasticity.

Table 4.3: Testing for a unit root and for homoscedasticity in the filtered series (p -values) when (4.1) is divided by unfiltered $\hat{\sigma}_t$ (no HP) and according to (4.4) (HP)

	US		UK		Aus	
	no HP	HP	no HP	HP	no HP	HP
ADF	0.00	0.00	0.00	0.00	0.00	0.00
P-P	0.00	0.00	0.00	0.00	0.00	0.00
ARCH-LM	0.20	0.71	0.66	0.90	0.81	0.63
B-P	0.03	0.31	0.43	0.89	0.84	0.93

As expected, the null hypotheses of a unit root is rejected in all cases. Also, filter (4.4) successfully removes heteroscedasticity from the US, UK and Aus GDP series, without seriously affecting the dynamics of the series. Note that the Breusch-Pagan (B-P) test signals heteroscedasticity (at significance level 0.05) in the filtered US series when no HP filter has been used to smooth the volatility.

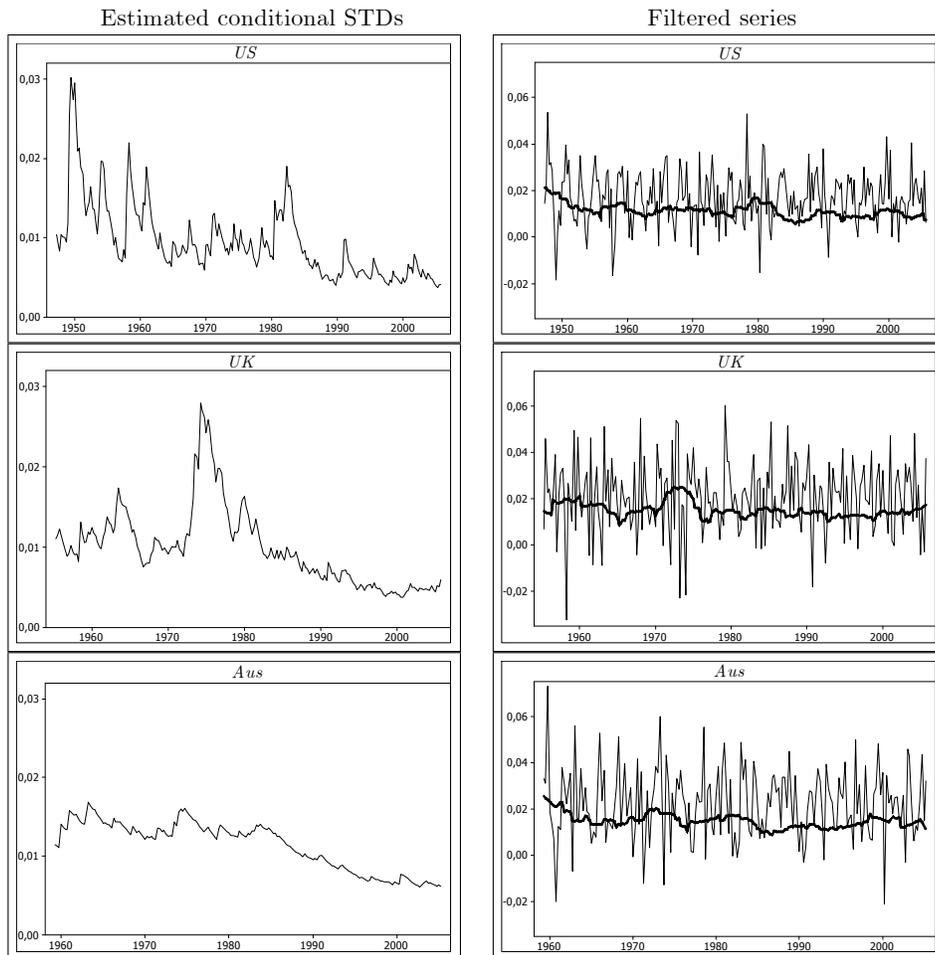
4.2 A GARCH-type filter

It is interesting to compare these results to the other approach mentioned earlier, namely to divide (4.1) by the conditional volatility estimated from GARCH-type models. Specifically, in the presence of asymmetries and leptokurticity, non-linear GARCH models must be considered which have the ability to capture asymmetric effects. Several such models exist in the literature, most notably the EGARCH model of Nelson (1991), the TGARCH model of Zakoian (1994) and the GJR-GARCH model of Glosten et. al (1993). These models were originally introduced to capture the leverage effect of stock returns. The models that are found adequate and minimize Akaike's criterion are:

$$\begin{aligned}
 z_{t,US}^{(a)} &= \varepsilon_t \\
 \ln(h_{t,US}) &= \underset{(0.26)}{-0.66} - \underset{(0.14)}{0.43} \left(\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \underset{(0.11)}{0.42} \left(\left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| - E \left(\left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| \right) \right) + \underset{(0.03)}{0.93} \ln(h_{t-1}) \\
 z_{t,UK}^{(a)} &= \underset{(0.07)}{0.22} z_{t-1,UK}^{(a)} + \varepsilon_t \\
 \ln(h_{t,UK}) &= \underset{(0.08)}{-0.15} - \underset{(0.02)}{0.11} \left(\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right) + \underset{(0.09)}{0.30} \left(\left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| - E \left(\left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| \right) \right) + \underset{(0.02)}{0.98} \ln(h_{t-1}) \\
 z_{t,Aus}^{(a)} &= \varepsilon_t \\
 h_{t,Aus} &= \underset{(0.03)}{0.06} \varepsilon_{t-1}^2 + \underset{(0.04)}{0.93} h_{t-1}
 \end{aligned}$$

Note that the coefficient estimates of h_{t-1} are close to one indicating a strong persistence in the conditional variance, or outright nonstationarity. This is also seen in Figure 4.6 showing the conditional volatility estimated using the above models (left panel) and the corresponding filtered series (right panel).

Figure 4.6: EGARCH/GARCH filtered series

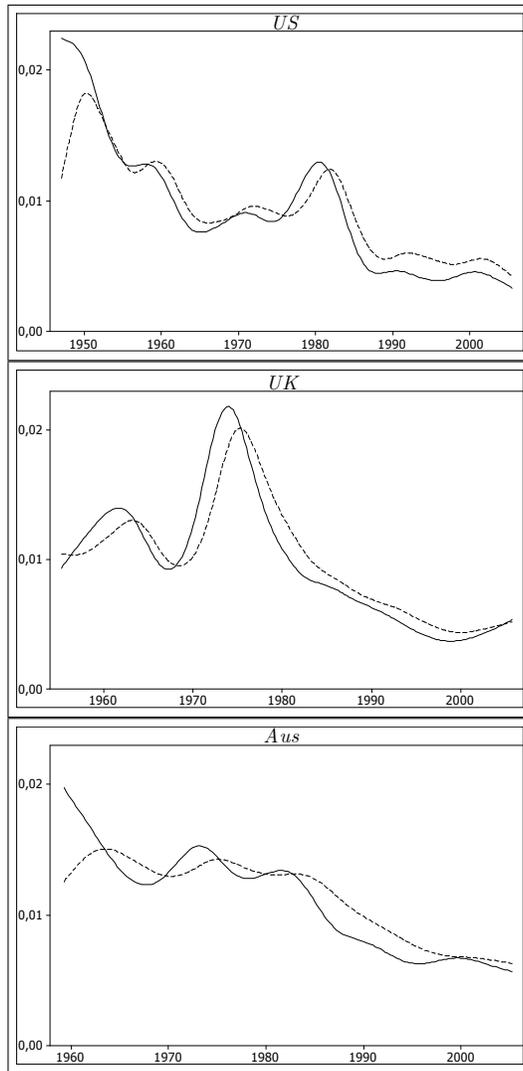


Note the rugged shape of the estimated volatility. It seems that HP-filtering of the volatility series would be even more needed here than in filter (4.4), see Figure 4.3. This is performed in the following subsection showing a comparison of the above filters.

4.3 Comparing the filters

Figure 4.7 shows a comparison of the HP trends (with $\gamma = 1\,600$) of the estimated volatility using filter (4.4) and the EGARCH/GARCH filter.

Figure 4.7: The HP trend of moving STDs of z_t for filter (solid lines) and of EGARCH/GARCH estimated STDs (dashed lines)



The amplitude of the HP trend is higher using filter (4.4) than the alternatives, rendering a heavier heteroscedasticity filtering. Also there are phase shifts between the HP trends, which is due to the fact that filter (4.4) is centered and the

EGARCH/GARCH is a causal filter. This can be corrected for by lagging the HP trend according to the calculated phase shift or constructing a two-sided filter, as in model based seasonal adjustment, see Hillmer and Tiao (1982). Here, the maximum value of the phase spectrum is 12.3, 10.7 and 10.9 quarters for the US, UK and Aus respectively. After correcting the series, almost identical filters are obtained. Note however that this procedure results in losing observations in the end of each series.

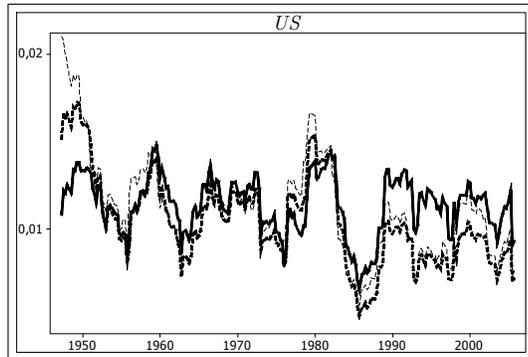
Despite the rather unstable looking moving STDs of the EGARCH/GARCH filtered series, the ARCH-LM tests fail to reject the null hypotheses of homoscedasticity, see Table 4.4

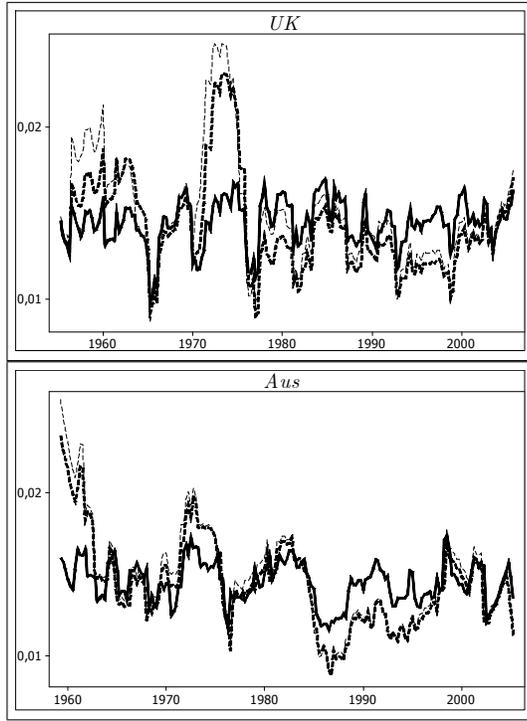
Table 4.4: Test for a unit root and homoscedasticity in the EGARCH/GARCH filtered series (p-values)

	US		UK		Aus	
	no HP	HP	no HP	HP	no HP	HP
ADF	0.00	0.00	0.00	0.00	0.00	0.00
P-P	0.00	0.00	0.00	0.00	0.00	0.00
ARCH-LM	0.97	0.31	0.39	0.68	0.65	0.80
B-P	0.90	0.11	0.07	0.32	0.44	0.83

Whether the HP filter renders the heteroscedasticity filtering more effective or not is still not clear. The p-values for the UK and Aus are higher using the HP filter than without, but it is the other way around for the US. Note, however, that the p-values for the ARCH-LM and B-P tests of the filtered series are considerably smaller on average compared to the ones in Table 4.3. To elucidate the differences, Figure 4.8 presents the moving STDs (again using $k = 15$) of each filtered series. The \tilde{z}_t series without the HP filter have been omitted in order to simplify comparisons. Also, to enable graphical comparisons in both ends, filter (4.4) has been modified there.

Figure 4.8: The moving STDs using filter (4.4) (solid line), and the EGARCH/GARCH filter (dashed line), with HP (thick line) and without HP (thin line)

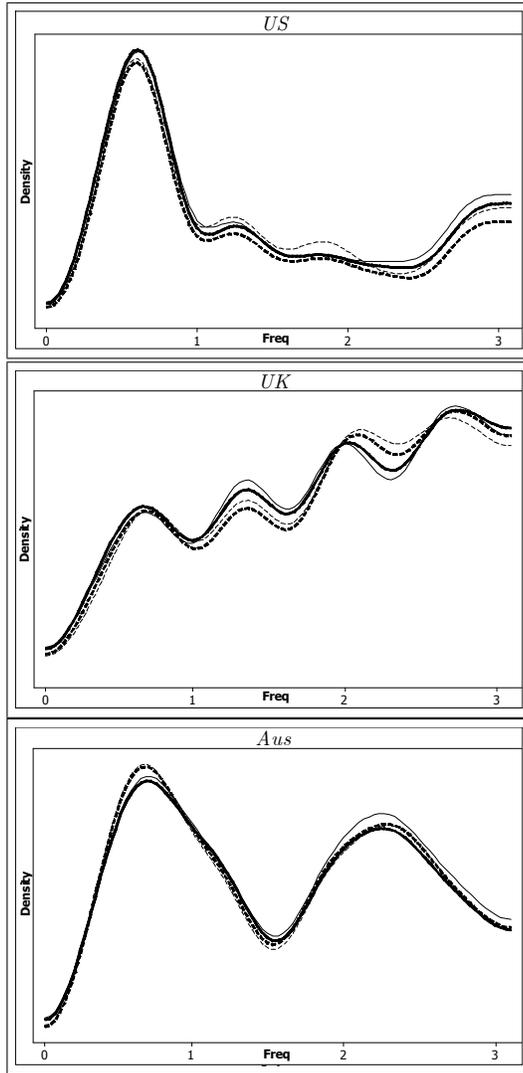




For all three countries it is again evident that filter (4.4) more effectively removes heteroscedasticity in these series. The moving STDs of the EGARCH/GARCH filtered series all contain a decreasing local trend in the beginning, and after that the amplitude of the swings seem to be larger. This is especially true when no HP filter was used in the filtering. This result is somewhat surprising, particularly for the US for which both the ARCH-LM and B-P tests in Table 4.4 report very high p-values compared to the other filters. According to e.g. the sign test, the median absolute deviations from the overall STD is significantly smaller using filter (4.4) than the EGARCH/GARCH filter. Applying the HP filter on the STDs or not does not change this result. Our recommendations to apply it are based on logical considerations, see the discussion in Section 4.1. Also, this way of reasoning is supported by the test results in Table 4.3, 4.4, the sign test and Figure 4.8.

Figure 4.9 shows that the spectral densities of the filtered series using (4.4) or the EGARCH/GARCH filter, with and without the HP filter, are almost identical.

Figure 4.9: Spectral densities of the filtered series using (4.4) (solid line), and the EGARCH/GARCH filter (dashed line), with HP (thick line) and without (thin line)



The close similarities of the spectral densities are to some extent due to the Parzen window and truncation point used, $M = 20$, which for each country is between the existing rules-of-thumb values $M = \sqrt{n}$ and $M = 2\sqrt{n}$ suggested in the literature. By reducing M , larger differences will appear in the spectral densities, but the fundamental features will still be similar. An application of the simple filter (4.4) is the subject of the next section.

5. A forecast competition

In this section we generate 1-20 quarters ex ante forecasts for the original series, using adequate ARIMA models estimated for the proposed filtered series $\tilde{z}_{t,US}$, $\tilde{z}_{t,UK}$ and $\tilde{z}_{t,Aus}$. In table 5.1 the results for *filtered* data are marked with an asterisk *. We compare the accuracy of these with forecasts from ARIMA, ARFIMA and GARCH models for *unfiltered* (Diff ln) data. The accuracy of the out of sample forecasts is measured by the root mean square forecast error (RMSFE) divided by the STD, and multiplied by 100. Hence, RMSFE = 100 would signal that the forecast error is of the same size as that of a naive forecast. The fact that the values are considerably smaller, even for the longest horizons, is another symptom of the decreasing variance of the time series studied. For instance for Diff ln US GDP the STD of the latest 15 years is just 41 per cent of the STD of the entire period.

Table 5.1: GDP forecast comparisons⁴ (RMSFE)/STD

Country	ARIMA	ARIMA*	ARFIMA	GARCH	Horizon (q)
US	25.5	25.3	37.7	25.6	4
UK	51.2	40.8	55.7	61.2	4
Aus	27.8	19.3	42.7	36.7	4
Country	ARIMA	ARIMA*	ARFIMA	GARCH	Horizon (q)
US	24.6	36.4	27.2	36.0	8
UK	47.4	38.3	49.1	46.5	8
Aus	50.2	42.1	40.0	46.7	8
Country	ARIMA	ARIMA*	ARFIMA	GARCH	Horizon (q)
US	46.3	45.9	48.6	41.7	12
UK	39.8	36.7	40.2	38.2	12
Aus	53.7	50.8	44.9	50.2	12
Country	ARIMA	ARIMA*	ARFIMA	GARCH	Horizon (q)
US	80.5	69.8	71.5	97.9	20
UK	36.9	36.3	46.2	98.2	20
Aus	50.6	45.5	42.3	95.4	20

Bold figures mark the lowest RMSFE of each row. Table 5.1 offers little support to Candelon and Gil-Alana (2004); ARFIMA models produce the least accurate forecasts at the max. one year horizon. For the longer horizons ARFIMA has the lowest RMSFE only for Australian GDP. Estimating ARIMA models from the filtered series, rather than from the unfiltered ones, generates significantly more accurate forecasts when pooling across all horizons above, according to

⁴The forecasts with a horizon of four quarters are made for 2005, using data up to quarter 4 of year 2004 for estimation. Consequently, four point forecasts are compared to four actual outcomes. Similarly, the outcomes of 2004-2005, 2003-2005 and 2001-2005, are omitted in the estimation step and used in the comparisons of the 8, 12 and 20 quarter forecast horizon respectively.

the Diebold-Mariano test of equal forecasting performance (p value = 0.02). When comparing our method with ARFIMA and GARCH-type models (also estimated on the unfiltered series) the p values of the Diebold-Mariano test is 0.04 and 0.01, respectively. The superiority of models based on filtered data is due to the inefficiency of estimates from heteroscedastic data. The series in this study can be seen as a worst case scenario, because the volatility of the series decreases over time. Consequently, parameter estimates are strongly based on an obsolete structure.

6. Conclusions

In this paper we discussed the issue of removing heteroscedasticity. We also propose a simple filter that successfully removes the heteroscedasticity in GDP series without significantly distorting the dynamics. Unlike the Box-Cox transformation, the filter does not assume that the heteroscedasticity is proportional to the level of the series. Dividing the heteroscedastic series by ARCH/GARCH estimated volatility is much more cumbersome and is significantly less effective than the proposed filter.

A mechanical estimation of e.g. ARIMA or ARFIMA models on heteroscedastic Diff ln GDP series is unbiased, but inefficient. Using filtered data resulted in better forecasts in a large majority of cases. When pooling across all horizons, the estimated ARIMA models from the filtered series generated significantly more accurate forecasts compared with ARIMA, ARFIMA and GARCH on the unfiltered data. The result is based on only three series, but they are the longest and the most important quarterly GDP series around. This simple filter could be used as a standard method to remove heteroscedasticity, much as seasonality is suppressed by seasonal adjustment filters.

Acknowledgments: We gratefully acknowledge helpful comments from Daniel Thorburn of Stockholm University, Mattias Villani of the Swedish Riksbank and from seminar participants at Cass Business School (London) Stockholm University, University of Helsinki, Åbo Akademi University (Turku) and the International Symposium on Forecasting in New York and in Nice. Research was partially supported by the International Institute of Forecasters, the Royal Swedish Academy of Sciences and by Societas Scientiarum Fennica.

References

- Baltagi, B. H., Jung, B. C. and Song, S. H. (2009) Testing for heteroskedasticity and serial correlation in a random effects panel data model. *Journal of Econometrics*, 154, 122-124.
- Box, G. E. P. and Cox, D. R. (1964) An analysis of transformations. *Journal of Royal Statistical Society*, B, 26, 211–252
- Candelon, B. and Gil-Alana, L. A. (2004) Fractional integration and business cycle features. *Empirical Economics*, 60, 343-359.
- Cavaliere, G. (2004) Unit root tests under time-varying variances. *Econometric Reviews*, 23, 259-292.
- Giordani, P. and Villani, M. (2010) Forecasting macroeconomic time series with locally adaptive signal extraction. *International Journal of Forecasting*, Forthcoming.
- Glosten, L., Jagannathan, R. and Runkle, D. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5), 1779-1801.
- Hamori, S. and Tokihisa, A. (1997) Testing for a unit root in presence of a variance shift. *Economics Letters*, 57, 245-253.
- Hess, G. D. and Iwata, S. (1997) Measuring and comparing business-cycle features. *Journal of Business and Economics and Statistics*, 15, 432-444.
- Hillmer, S.C. and Tiao, G.C. (1982) An ARIMA-model based approach to seasonal adjustment. *Journal of the American Statistical Association*, 77, 63-70.
- Hodrick, R. J. and Prescott, E. C. (1997) Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29, 1-16.
- Kim, K. and Schmidt, P. (1993) Unit root tests with conditional heteroskedasticity. *Journal of Econometrics*, 59, 287–300.
- Kim, T. H., Leybourne, S. and Newbold, P. (2002) Unit root tests with a break in innovation variance. *Journal of Econometrics*, 109, 365-387.
- Nelson, B. (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59, 347-370.
- Wallis, K. F. (1974) Seasonal adjustments and relations between variables. *Journal of the American Statistical Association*, 69, 18-31.
- Zakoian, M. (1994) Threshold heteroscedastic models. *Journal of Economic Dynamics and Control*, 18, 931-955.