RESEARCH REPORT SERIES
*(Statistics #2019-02)*

# A class of multivariate filters for trend extraction and statistical analysis of multiple related time series

Thomas M. Trimbur,
Tucker S. McElroy

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: February 25, 2019

# A class of multivariate filters for trend extraction and statistical analysis of multiple related time series

Thomas Trimbur and Tucker McElroy

U.S. Census Bureau

**Abstract**

This article develops a class of multivariate filters for extracting related signals from multiple nonstationary time series affected by noisy fluctuations. Many such applications are possible for the statistical data available in various sciences. For instance, in the bivariate case, the filters can be used for signal analysis where one series is of prime interest and is related to a second series with higher signal-content; the formulas show exactly how much weight should be placed on the auxiliary data at different leads and lags. The multivariate class generalizes the widely used Butterworth class, which has the limitation of being applicable only separately to each individual series, even in multivariate applications. The new filters provide the same flexibility in design – with minimal complexity of form – as basic Butterworth filters. In the multiple series case, there are similarly compact gain functions that now account for inter-relationships among series. The filter parameters, which may have important effects on conclusions based on the extracted signals, may be guided by the dataset at hand to help ensure consistency with observed properties. An application to U.S. petroleum consumption is presented, where more precise trends are estimated by making use of the related time series of OPEC oil imports.

**Keywords.** Co-integration; Common Trends; Filters; Multivariate Models; Stochastic Trends; Unobserved Components.

**Disclaimer** This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau.

# 1 Introduction

A great deal of the time series data encountered in the social and physical sciences has a stochastic element, with unpredictable and erratic noise superimposed on underlying signals or regular movements that represent the object of study. Filters can be applied to a time series, with weights attached to observations at different leads and lags, to smooth out the data and estimate such signals by removing any noisy or extraneous components. The estimation of trends is a pervasive area of interest and represents a key part of much research in the sciences, for instance being extensively used in economics and climatology. Applications in economics include Dimitropoulos et al (2005), Agnolucci (2010), Adeyemi et al (2010), Broadstock et al (2011), Hunt et al (2003); while trends in climatological data were examined by Gallant et al (2014), Coumou et al (2013), and Visser et al (2015, 2018).

In the literature on signal extraction theory and methodology, Wiener (1949) and Whittle (1963) made substantial early contributions; the Wiener-Kolmogorov (WK) formula gives the asymptotic form (for historical, or two-sided smoothing of doubly infinite series) of the relationship between optimal signal estimation and component properties. Bell (1984) extended this formula to the important nonstationary case. Gómez (2001) and Harvey and Trimbur (2003) explored the connection between unobserved components (UC) models, which decompose a series into latent processes such as stochastic trends, and filtering methods that weight observed data at various leads and lags.

"Butterworth filters", so-called because they were introduced in Butterworth's (1930) research in electrical engineering, are commonly used to extract signals. Over the past several decades, Butterworth filters have played a role in numerous applications in fields such as astronomy, economics, medicine, atmospheric science, climatology, and oceanography. For some examples, see Gao et al (2012), Avdeeva et al (2014), Gehring et al (2013), and Tauzin et al (2010). The popularity of this class of filters owes to their combination of an efficient representation and compact form in the frequency domain, together with a flexibility in effects and gain properties that can be easily and transparently controlled by the researcher. Gómez (2001) showed that Butterworth low-pass filters are directly linked to standard UC models containing trend and noise components. Their gain function, which shows how they affect the various component frequencies of an observed series, allows for direct choice of cutoff frequency and sharpness as filter parameters.

Standard Butterworth filters are applied separately to each individual series. However, in current practice and research, where available, related variables are often combined to attain more informed assessments of a scientific phenomenon or regular dynamic. Further, most of the data encountered in the social and physical sciences has a stochastic element, with the background noise leading to uncertainty in the estimated signal, and many time series have nonstationary movements. Hence, there is a need for extending Butterworth filters to the multivariate setup, in a way that

1) the filters make efficient use of series' interrelationships, 2) the uncertainty associated with the filtering is considered, and 3) the filters handle nonstationarity and inter-relationships among nonstationary movements and among stationary parts.

To attain these objectives, this paper develops a generalized class of multivariate Butterworth filters for isolating trends and other signals in related, multiple series. The filters have compact gain functions like basic Butterworth filters, and they account for dynamic relationships across (possibly nonstationary) series while overcoming limitations of the current methodology.

In previous work, to our knowledge the analysis of Butterworth filters and their application have been conducted only for the univariate case. The multivariate generalizations of Butterworth filters that we introduce make use of a (time series) model-based approach; generalizations of models such as those in Harvey and Koopman (1997) are considered. A prime motivation for the multivariate filters is to take advantage of correlations in the fluctuations of related variables; yet there is no clear way to extend a Butterworth filter used in nonparametric fashion in this direction. By analyzing the implicit statistical models, we can achieve the generalized filters, while at the same time providing the means for 1) incorporating cross-correlations into the filter design, 2) assessing the degree of uncertainty in target signal, 3) handling nonstationarity and special inter-relationships among the trends, 4) taking consideration of consistency with data properties in designing filters, and 5) adapting filters near series' endpoints.

Application of the methodology is presented for the statistical measurement of the trend in a time series of U.S. petroleum consumption. The dynamic representations in terms of trend plus noise models give an essential starting point for describing the behavior of the target consumption series and the ancillary time series of petroleum imports from the OPEC cartel. The multivariate formulas for the two-series case show the optimal combination of two series – with rather different properties – in the formation and application of estimated bivariate low-pass filters. The bivariate case already helps to give the main insights into the advances offered by multivariate relative to univariate filter design. The enhancements, such as more efficient use of data, inclusion of information about cross-relationships, improvements in modelling performance, reduction in measured signal uncertainty and signal growth rate uncertainty, are already illustrated in such applications. The bivariate extension is useful for developing intuition about filter formation in terms of weighting observations on a pair of time series and applying gain functions that include cross-relations.

While trends in Energy Consumption have been explored by numerous researchers – among others, Dimitropoulos et al (2005) and Agnolucci (2010) – our analysis represents the first use of OPEC imports data to improve on trend estimation for oil consumption. The OPEC imports series is a natural candidate for bivariate signal estimation, since it is observed with substantially less irregular movements than consumption and has trend movements that correlate with and lead the trend in consumption – a connection that is intuitive by virtue of OPEC's role of supplying the marginal barrel of oil. In detecting signals from these noisy datasets, we demonstrate the improvement of signal estimates and modelling aspects when using bivariate time series models as the basis of the signal estimation. The results for the estimated filters quantify the degree of emphasis on imports – the high signal-content series – and the corresponding reduced weighting on consumption itself, which can be expressed in either the time or frequency domains. The trend differs importantly from the one measured with a simple reliance on consumption alone and can be estimated with greater accuracy. A certain degree of emphasis falls on the series with richer signal content, OPEC imports, in removing the transitory and high-frequency parts. These results have implications for the modelling and assessment of trends in Oil Consumption more broadly, as the stochastic trend model we employ could be directly incorporated into structural equations relating Consumption to factors like product prices and income.

The filter introduced by Hodrick and Prescott (1997) has been popular in macroeconomic analysis; multivariate extensions were proposed by Dermoune et. al (2009) and Poloni and Sbrana (2017). Our filter class is more general along the two dimensions of trend order and allowance for slope damping. Further, we produce generalized Butterworth filters that can be mapped into the original intuitive form of the filters in terms of cutoff frequency and sharpness in the gain function.

The rest of the paper is arranged as follows. Section 2 reviews Butterworth filters and their time domain representation that may be less familiar to many readers. In Section 3, a simulated illustration is given to provide motivation for the general multivariate filters. Section 4 proceeds to set out models for multiple related time series, from which the multivariate class of Butterworth filters are derived. Then in Section 5, we apply the methodology to estimate the trend in petroleum consumption using the related signal of market conditions given by OPEC oil imports. Section 6 gives concluding remarks.

## 2 Butterworth Filters and the statistical connection

This Section reviews Butterworth filters and briefly notes their connections with statistical models expressed as unobserved components. This linkage was established in Gómez (2001) and Harvey and Trimbur (2003). In this paper, we will make use of the Butterworth filter's underlying models to generalize this class to the multivariate setting; here, the model-filter connection becomes crucial for providing guidance regarding the diverse set of possible generalizations and for making the best possible use of series' relationships in pinpointing their low-frequency components. Section 4 below introduces this generalized class of filters. The current Section reviews the univariate case.

## 2.1 Butterworth filters - original forms

The existing Butterworth filters are typically defined in terms of their gain function, which shows how different frequency components are attenuated when applying the filter. Letting $\lambda$ denote frequency in radians, the gain is expressed as

$$B_m^{lp}(\lambda) = \left[1 + \left(\frac{\sin(\lambda/2)}{\sin(\lambda_{lp}/2)}\right)^{2m}\right]^{-1}, \tag{1}$$

where $\lambda_{lp}$ is the frequency at which the gain equals one-half, and $m$ is a positive integer denoting the order, or index, of the filter. This form was originally presented in Butterworth (1930). The low-pass filter aims to cut out frequencies beyond $\lambda_{lp}$, with the higher frequencies cut off more sharply as $m$ increases. The gain function shows how they remove noisy fluctuations in data by cutting out higher frequencies, allowing researchers to focus on the more regular, low-frequency signals.

The class of Butterworth low-pass filters provide flexibility in location of cutoff frequency and sharpness of gain, and this flexibility, in addition to the simple form of the gain, has underpinned their widespread use in statistical analysis. Formula (1) refers to the "Butterworth-sine" form, which we will also refer to as the "standard" form, given its frequent appearance in the literature and its connection with commonly used models of stochastic trends.

Another class, which has a closely related form, is the general "Butterworth-tangent" filter, whose gain is

$$B_m^{\tan}(\lambda) = \left[1 + \left(\frac{\tan(\lambda/2)}{\tan(\lambda_{lp}/2)}\right)^{2m}\right]^{-1}. \tag{2}$$

This filter class, less frequently appearing in previous work, provides the same compactness of gain function, and its time domain representation will be explored alongside the standard class in this paper. We will also refer to this filter class as "canonical" Butterworth, given their close linkage with the canonical decomposition of time series, as originally introduced in Tiao and Hillmer (1978) and Hillmer and Tiao (1982). This decomposition is based on the fundamental and intuitive principle of extracting as much noise as possible (hence maximizing the variance of the irregular/noise process).

Hence, the standard/sine and canonical/tangent classes of Butterworth filters provide some complementarity in how trends and signals are pinpointed from fluctuating time series.

## 2.2 Time Domain Expressions for filters and implicit models

When a gain function – such as those noted above – is applied to a time series of observations $y_t$, the corresponding operation is to weight the observations according to

$$B_m^{lp}(L)y_t = \sum_j w_j y_{t-j} \tag{3}$$

which, using $L$ to denote the lag operator (so that $Ly_t = y_{t-1}$), can be written as

$$B_m^{lp}(L) = \sum_{h=-\infty}^{\infty} w_h L^h. \tag{4}$$

The set of weights, the values of $w_h$ at different leads and lags, act to smooth out fluctuations in the series. This produces a trend estimate whose weights are the inverse Fourier Transform of the frequency response (which is equivalent to the gain for symmetric filters with real weights). For the Butterworth filters, $B_m^{lp}(L)$ takes on compact expressions that are straightforward to derive, given the original gain functions.

Beginning with the standard Butterworth gain, using basic trigonometric identities, it can be shown that

$$B_m^{lp}(\lambda) = \frac{1}{1+\frac{1}{q}(2-2\cos\lambda)^m}, \qquad 0 \leq \lambda \leq \pi, \tag{5}$$

with

$$q = [2\sin(\lambda_{lp}/2)]^{2m}, \qquad 0 < \lambda_{lp} < \pi. \tag{6}$$

Therefore, working back from the frequency domain to the time domain gives

$$B_m^{lp}(L) = \frac{1}{1+\frac{1}{q}[(1-L)(1-L^{-1})]^m} \tag{7}$$

as the coinciding expression for the standard Butterworth filters.

Similarly, for the Butterworth tangent,

$$B_m^{\tan}(L) = \frac{[(1+L)(1+L^{-1})]^m}{[(1+L)(1+L^{-1})]^m + \frac{1}{q^\dagger}[(1-L)(1-L^{-1})]^m} \tag{8}$$

with

$$q^\dagger = [2\tan(\lambda_{lp}/2)]^{2m}, \qquad 0 < \lambda_{lp} < \pi$$

The time and lag indices are doubly infinite in (4), and this asymptotic case pertains to the exact, ideal gain functions in (1) and (2) and, correspondingly to the closed-form filter expressions in (7) and (8).

Though Butterworth filters were originally introduced in the context of electrical engineering, which involves deterministic and often band-delimited processes, since their introduction they have increasingly been used in the context of statistical data consisting of stochastic movements, where the goal is to remove noisy, predominantly random, fluctuations. Hence, it is intuitive that expressions (1) and (2) relate to the extraction of a trend – from a process that includes both trend and noise components. That is, when the observations $y_t$ comprise certain models for stochastic trend, $\mu_t$, and noise $\varepsilon_t$,

$$y_t = \mu_t + \varepsilon_t$$

a Butterworth filter is specifically designed to extract $\mu_t$ with the greatest accuracy (on average when taken over repeated samples).

Indeed, the $m-$th order standard Butterworth filter gives the optimal way, in the sense of Minimum Mean-Squared Error (MMSE), to filter the series when the trend follows

$$\mu_{m,t} = (1-L)^{-m}\zeta_t \tag{9}$$

and $\varepsilon_t$ is a serially uncorrelated process with mean zero, or simple White Noise denoted by $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. This result was noted in Gomez (2001), who made use of the nonstationary formula proven in Bell (1984). The class $\{\mu_{m,t}\}$ defines the most elementary nonstationary process for each integration order. It includes the random walk and the "Smooth Trend" model as specific cases. The "Smooth Trend" Model (STM), so called because it often produces visible smoothness in the extracted trend, results from setting $m = 2$; in this case, $\mu_{2,t}$ represents an I(2) process that accounts for a time-varying slope.

In a similar way, the $m-$th order Butterworth tangent filter is the optimal estimator of the following stochastic trend:

$$\mu_{m,t}^{\dagger} = (1+L)^m(1-L)^{-m}\zeta_t$$

when $\mu_t$ is given by $\mu_{m,t}^{\dagger}$; see Gomez (2001).

This close connection between filtering and statistical modelling has been little studied with the exception of works such as Gómez (2001) and Harvey and Trimbur (2003), though it represents an invaluable tool for theory, methodology, and practice. The next Section provides some additional motivation for model-based multivariate filtering by way of example.

# 3 Illustration of Multivariate Butterworth Filters

This Section gives a bivariate illustration to help motivate the use of more general multivariate filters. Two noisy time series, $y_{A,t}$ and $y_{B,t}$, are simulated with trends $\mu_{A,t}$ and $\mu_{B,t}$, respectively, that are closely related to each other. As shown in figure 1 by the solid black and dotted blue

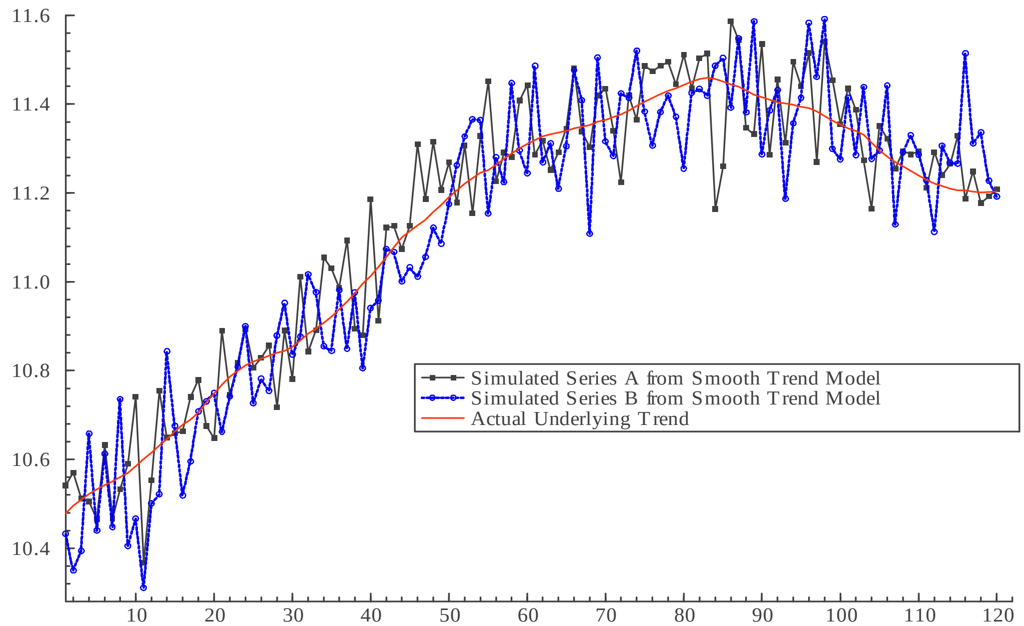


Figure 1: Two simulated noisy time series, $y_{A,t}$ and $y_{B,t}$, shown along with their common trend $\mu_{A,t}$ $= \mu_{B,t}$.

lines, though their general levels tend to track each other, both $y_{A,t}$ and $y_{B,t}$ are observed with high-frequency noise. The series were generated by the process:

$$\begin{bmatrix} y_{A,t} \\ y_{B,t} \end{bmatrix} = \begin{bmatrix} \mu_{A,t} \\ \mu_{B,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{A,t} \\ \varepsilon_{B,t} \end{bmatrix} \tag{10}$$

$$\varepsilon_{A,t}, \varepsilon_{B,t} \sim WN(0, \sigma_\varepsilon^2), \quad \text{E}[\varepsilon_{A,s}\varepsilon_{B,t}] = 0 \text{ for all s,t,}$$

where $\varepsilon_{A,t}$ and $\varepsilon_{B,t}$ are the irregular movements, or noise, in $y_{A,t}$ and $y_{B,t}$, respectively. The uncorrelatedness of the noise, an assumption that is relaxed in later Sections, simplifies the example here. Also, $\varepsilon_{A,t}, \varepsilon_{B,t}$ are assumed to be normally distributed and to have the same variance for both series. The data were simulated for $T = 200$ observations, with starting values $\mu_{A,0} = \mu_{B,0} = 10$; Figure 1 displays the middle 120 observations, which helps make the cross-relationships of the series clearer.

The trends follow a STM, that is, are given by the bivariate extension of (9) with $m = 2:$

$$\begin{bmatrix} \mu_{A,t} \\ \mu_{B,t} \end{bmatrix} = \begin{bmatrix} \mu_{A,t-1} \\ \mu_{B,t-1} \end{bmatrix} + \begin{bmatrix} \beta_{A,t-1} \\ \beta_{B,t-1} \end{bmatrix}, \tag{11}$$

$$\begin{bmatrix} \beta_{A,t} \\ \beta_{B,t} \end{bmatrix} = \begin{bmatrix} \beta_{A,t-1} \\ \beta_{B,t-1} \end{bmatrix} + \begin{bmatrix} \zeta_{A,t} \\ \zeta_{B,t} \end{bmatrix},$$

$$\zeta_{A,t}, \zeta_{B,t} \sim WN(0, \sigma_\zeta^2), \quad \text{Corr}[\zeta_{A,t}\zeta_{B,t}] = \nu_\zeta \text{ for all t}, \quad \text{E}[\zeta_{A,s}\zeta_{B,t}] = 0 \text{ for all s,t with s} \neq t,$$

so that the trend in series $y_{A,t}$ is incremented by $\beta_{A,t-1}$ each period, and likewise for $y_{B,t}$. Therefore, $\beta_{A,t}$ and $\beta_{B,t}$ are slope processes that represent the rate of change in the trends; here, they are assumed to be random walk processes driven by correlated white noise disturbances $\zeta_{A,t}$ and $\zeta_{B,t}$ with the same variance. They are initialized as $\beta_{A,0} = \beta_{B,0} = 0$, so the trends have initial values $\mu_{A,1} = \mu_{B,1} = 10$ for the first pair of observations at $t = 1$.

For this illustration, it was further assumed that perfect correlation holds, that is $\nu_\zeta = 1$. As the trends have a linear relationship with the coefficient set to unity and the intercept set to zero, they can be expressed in terms of a single underlying trend (displayed in Figure 1 in red):

$$\begin{bmatrix} y_{A,t} \\ y_{B,t} \end{bmatrix} = \iota_2 \mu_t + \begin{bmatrix} \varepsilon_{A,t} \\ \varepsilon_{B,t} \end{bmatrix} \tag{12}$$
$$\mu_t = \mu_{t-1} + \beta_{t-1}$$
$$\beta_t = \beta_{t-1} + \zeta_t$$
$$\zeta_t \sim WN(0, \sigma_\zeta^2),$$

where $\iota_2$ is a 2 x 1 column vector of ones. The processes $y_{A,t}$ and $y_{B,t}$ start out with a flat underlying growth-rate that increases gradually over time, so that as indicated in figure 1, by $t = 40$ the trend is increasing steadily.

Assume that a researcher working with the dataset $\{[y_{A,t}, y_{B,t}], t = 1, ..., T\}$ wishes to measure the trend in $y_{A,t}$. As a first case, suppose that they have knowledge of the Data Generating Process (DGP) for $y_{A,t}$, that is the univariate form of the smooth trend plus noise model, given by

$$y_{A,t} = \mu_{A,t} + \varepsilon_{A,t},$$

with trend defined as the first row of (11). Further, while they are assumed to also know the values $\{\sigma_\varepsilon^2, \sigma_\zeta^2\}$, they use univariate filters applied to $y_{A,t}$ and omit any information about $y_{B,t}$ in the measurement of $\mu_{A,t}$; this omission could result from being constrained to simple univariate filters or from lack of knowledge about the relations between $y_{A,t}$ and $y_{B,t}$.

Then, the optimal (asymptotic) filter for extracting $\mu_{A,t}$ is given by (7) with $m = 2$ and $q = 10^{-3}$, and the corresponding gain function has the form (1), where $\lambda_{lp} = \lambda_{lp}^*$ is found by solving (6) for $\lambda_{lp}$. As illustrated in Figure 2, the gain function passes low frequencies and tapers off gradually to zero at higher frequencies. Given the dataset, using this standard Butterworth filter and applying it to $y_{A,t}$ to estimate $\mu_{A,t}$ yields the estimated trend indicated in Figure 4 (The standard Butterworth and split gain Butterworth filters can be applied using Kalman filtering).

In extending the univariate filter to extract the trend in series A, a filter can be applied to each series and the resulting output summed together. Therefore, as a second case, to obtain the estimate $\mu_{A,t}$, the researcher applies an "auto-filter" to the same series, $y_{A,t}$, along with a "cross-filter" to

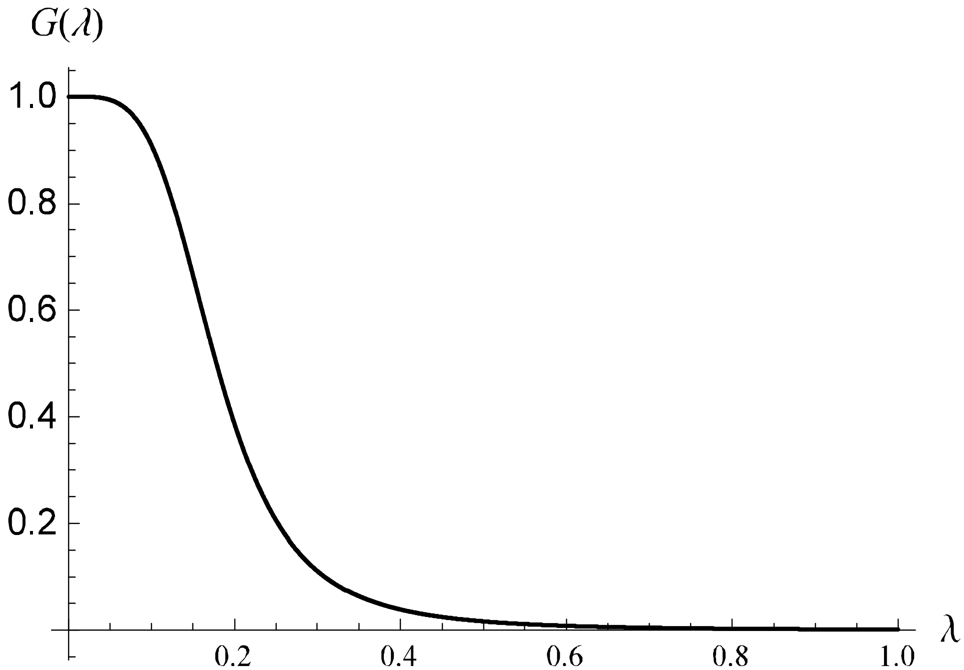


Figure 2: Gain function of univariate low-pass filter of the Butterworth-sine class with $m = 2$ and $q = 10^{-3}$.

the other series, $y_{B,t}$, and then sums their respective contributions. An especially simple bivariate filter, and the one suggested by the symmetric properties and shared trend of the series, would involve equivalent gains for each series, given by one-half the univariate gain in (1). This "splitting of gains" is illustrated in figure 3, where each bivariate gain is given by the solid line, compared to the standard univariate gain given by the dotted line. The resulting trend in this case is compared to the univariate filter output and to the actual trend in figure 4.

The corresponding errors in these estimates are plotted in figure 5. Taking the simple average of the squared trend estimation errors for the univariate filter (over the middle 120 observations, as indicated in the figure, which helps abstract from near end-point effects) gives 0.000833, while the split-gain approach yields an average squared-error of 0.000460, a 44.8% reduction. The estimate of the square root of the Mean-Squared Error, or RMSE, based on the simulated sample falls 25.7% going from the univariate ($\widehat{RMSE} = 0.0289$) to the to the bivariate ($\widehat{RMSE} = 0.0214$) strategy.

With knowledge of the close trend relationship and of shared trend-noise properties, the use of symmetrically split-gain functions is intuitive here. In other cases of perfectly correlated, or co-integrated trends, the same principle of gain splitting applies with more general patterns for the coefficients for the auto- and cross-gains. More generally, $|\nu_\eta| \leq 1$, and trend and noise characteristics vary across different time series; at the outset, it is unclear how, exactly, to construct suitable multivariate filters to benefit from the combined information in multivariate datasets.

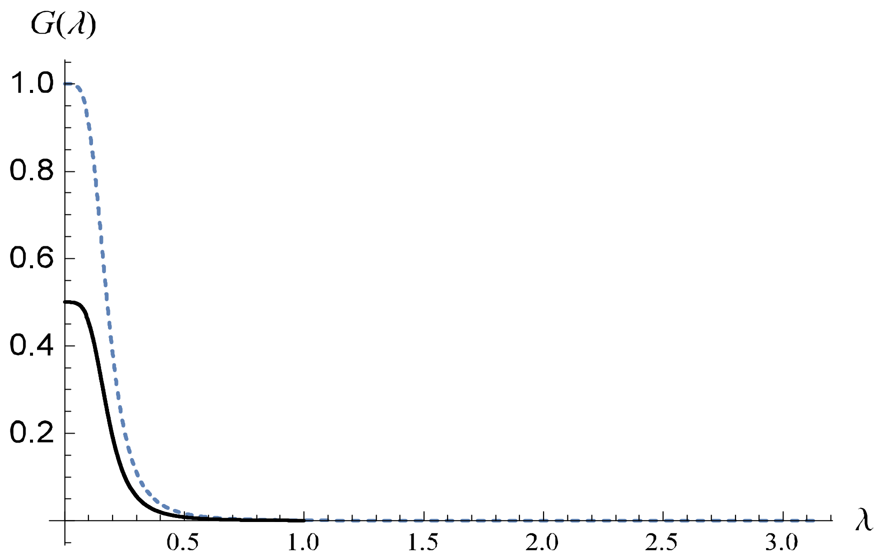


Figure 3: Gain function of bivariate low-pass filter based on Butterworth-sine class, with $m = 2$ , $q = 10^{-3}$, and $\nu_\zeta = 1$. The gain applied to each series, prior to summing their output, is indicated by the solid line. The univariate gain is displayed as the dotted line.

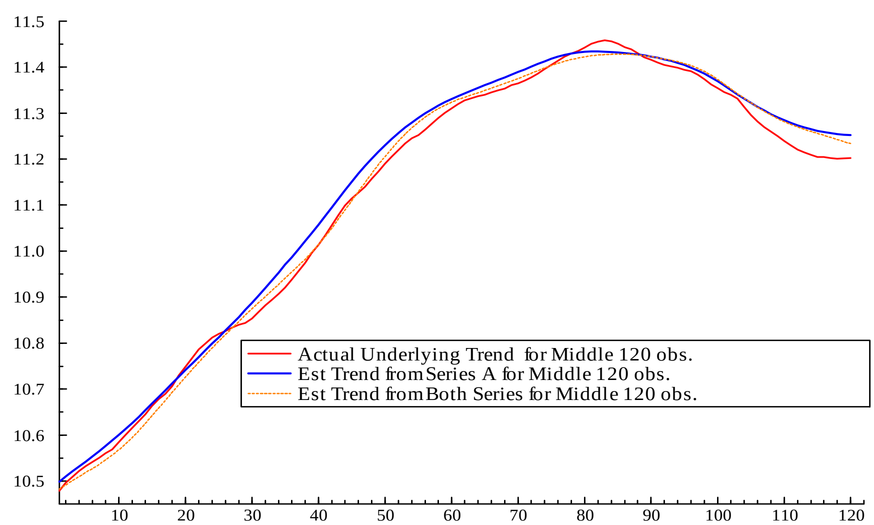


Figure 4: Trend estimates from the univariate filter applied to the series $y_{A,t}$ and from the bivariate filter applied to both series $y_{A,t}$ and $y_{B,t}$, compared to actual trend, shown for the middle 120 observations.

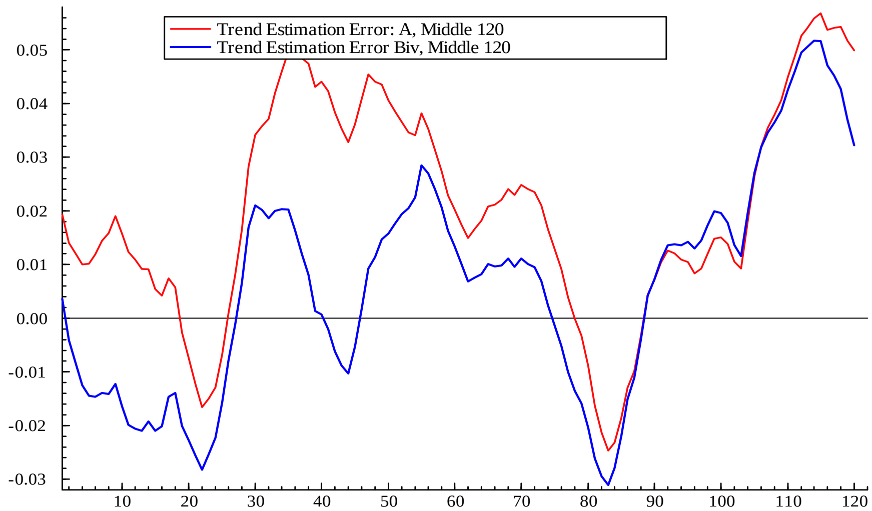


Figure 5: Errors in the trend estimates for the univariate filter applied to the series $y_{A,t}$ and for the bivariate filter applied to both series $y_{A,t}$ and $y_{B,t}$, shown for the middle 120 observations.

Hence, the generalization of the Butterworth filters involves some non-obvious aspects to provide the best statistical estimates of trends or correspondingly, the optimal elimination of high-frequencies, when series are related in flexible ways, and the researcher aims to judiciously combine information and reduce filtering error.

Finally, as a third case, suppose the researcher, in pursuit of greater smoothness in the estimated signal, applies a more stringent gain function to eliminate additional frequency components away from the low-end. This gain function, based on an order $m = 4$ Butterworth-sine filter with $q$ set to $\tilde{q} = 5 \times 10^{-9}$, is displayed in Figure 6 as the dotted line; the cutoff frequency is well below that for the optimal filter, and additionally the gain is sharper, or more block-like. Now, the average squared estimation error is 0.00099698, which is more than twice the value of 0.000460 obtained for the adaptive case. Hence, the pursuit of smoothness in the extracted trend should be moderated by consideration of the dynamic properties of the series.

# 4 Multivariate Butterworth filters

In this Section, we derive explicit expressions for the multivariate generalization of Butterworth low-pass filters based on the classes of m-th order trends, of both standard and canonical form. The Section provides the general formulation involving a $N$-dimensional time series, where $N$ is any positive integer, and also discusses more flexible trend structures. This treatment constitutes

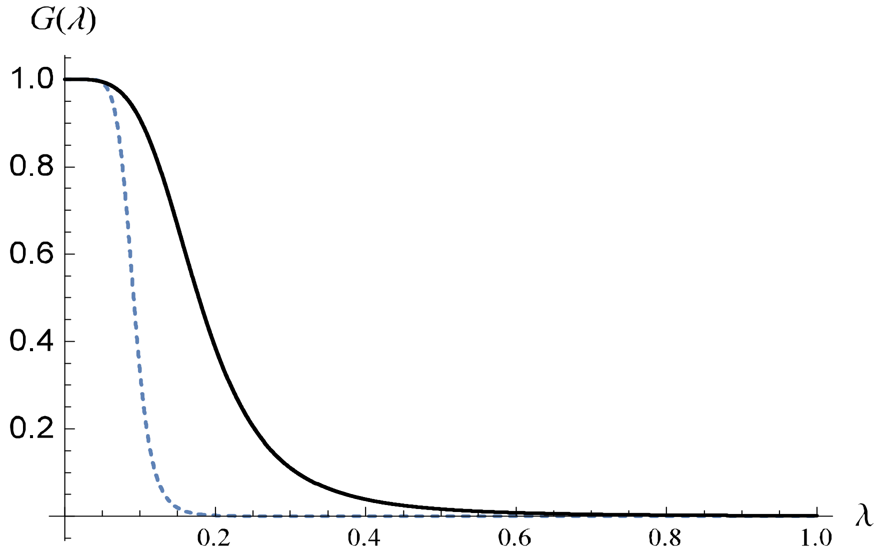


Figure 6: Gain functions of univariate low-pass filters of the Butterworth-sine class, first for $m = 2$ and $q = 10^{-3}$ and second, for $m = 4$ and $q = \widetilde{q} = 5 \times 10^{-9}$. Note the range of the x-axis is [0,1], not [0,$\pi$] as in figure 3.

the multivariate generalization of the Butterworth sine and tangent filters, whose univariate forms are reviewed in Gómez (2001) and Harvey and Trimbur (2003).

## 4.1 Models of related and Common Trends

Multivariate models capture the crucial aspect that related series undergo similar movements, and accounting for these cross-relationships can help give a better description of the fluctuations seen by each series. Models with related trends, where the underlying permanent shocks are correlated, allow us to establish links between series in their long-run behavior. This Section describes the general class of stochastic trend models for multiple time series, used to develop corresponding signal extraction filters by making use of the theory in McElroy and Trimbur (2015).

Let $N$ denote the number of series in the dataset. Define the vector process $\mathbf{y}_t = (y_t^{(1)}, ..., y_t^{(N)})'$ as the observed series at time $t$. Then the univariate trend-noise form directly generalizes to the following multivariate model:

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim WN(0, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}), \quad t = 1, ..., T \tag{13}$$

where $\boldsymbol{\mu}_t = (\mu_t^{(1)}, ..., \mu_t^{(N)})'$ is a stochastic trend vector and $\boldsymbol{\varepsilon}_t = (\varepsilon_t^{(1)}, ..., \boldsymbol{\varepsilon}_t^{(N)})'$ a vector of irregulars, both taken at time $t$.

Analogously to (9), in the multivariate case, the $m-th$ order standard trend is given by

$$\begin{aligned} \boldsymbol{\mu}_{i,t} &= \boldsymbol{\mu}_{i,t-1} + \boldsymbol{\mu}_{i-1,t-1}, \quad i = 2, ..., m \\ \boldsymbol{\mu}_{1,t} &= \boldsymbol{\mu}_{1,t-1} + \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \sim NID(0, \boldsymbol{\Sigma}_\zeta), \end{aligned} \tag{14}$$

so that at time $t$, the trend $\boldsymbol{\mu}_{m,t}$ is subject to a vector of slopes $\boldsymbol{\mu}_{m-1,t-1}$ that are themselves integrated processes for $m = 2$ and higher. The core disturbance vector is specified by its covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$, which is an $N \times N$ positive semi-definite matrix. The relations among the trends are reflected in the cross-correlations implied by $\boldsymbol{\Sigma}_\zeta$. For $m = 1$, only the equation for $\boldsymbol{\mu}_{1,t}$ is defined, giving a multivariate random walk as the stochastic trend. Many datasets can already be reasonably modeled with a once- or twice-integrated trend. Using $m > 2$ can also sometimes give viable specifications, and can help track steady variation in rate of change or in acceleration; such higher order trends give rise to sharper low-pass filters – which can produce attractive signal contours when applied to time series data, and this case should not necessarily be ruled out a priori.

For reduced rank $\boldsymbol{\Sigma}_\zeta$, the trend vector may be expressed as

$$\begin{aligned} \boldsymbol{\mu}_{m,t} &= \boldsymbol{\Theta}\boldsymbol{\mu}^\dagger_{m,t} + \mathbf{f}(\boldsymbol{\mu}_{1,0}, \boldsymbol{\mu}_{2,0}, ..., \boldsymbol{\mu}_{m-1,0}; t) \\ \boldsymbol{\mu}^\dagger_{i,t} &= \boldsymbol{\mu}^\dagger_{i,t-1} + \boldsymbol{\mu}^\dagger_{i-1,t-1}, \quad i = 2, ..., m \\ \boldsymbol{\mu}^\dagger_{1,t} &= \boldsymbol{\mu}^\dagger_{1,t-1} + \boldsymbol{\zeta}^\dagger_{1,t}, \qquad \boldsymbol{\zeta}_t \sim NID(0, \boldsymbol{\Sigma}_{\zeta\dagger}), \end{aligned} \tag{15}$$

where $\boldsymbol{\Theta}$ is an $N \times K$ matrix, with $K < N$. For the length-$N$ vector $\mathbf{f}(\boldsymbol{\mu}_{1,0}, \boldsymbol{\mu}_{2,0}, ..., \boldsymbol{\mu}_{m-1,0}; t)$, the first $K$ positions consist of zeroes , while the remaining $N - K$ elements contain deterministic polynomials in $t$, whose coefficients depend on the initial conditions for the trend process (including both the trend itself and any lower order processes used in defining the trend when $m > 1$). For identification, the elements of the load matrix $\boldsymbol{\Theta}$ are constrained to satisfy $\boldsymbol{\Theta}^{jk} = 0$ for $k > j$, and $\boldsymbol{\Theta}^{jj} = 1$ for $j = 1, ..., K$. The long-run movements then depend on the smaller set of processes, arranged in the $K$-element vector $\boldsymbol{\mu}^\dagger_t$ and driven by disturbance $\boldsymbol{\zeta}^\dagger_t$ with diagonal $(K \times K)$ covariance matrix, $\boldsymbol{\Sigma}_{\zeta\dagger}$. For these identifying conditions on $\boldsymbol{\Theta}$ and $\boldsymbol{\Sigma}_{\zeta\dagger}$, the "common trends" in $\boldsymbol{\mu}^\dagger_t$ have the interpretation of distinct or elemental long-term patterns that are "pure" in the sense of being completely unrelated to each other. McElroy (2017) also discusses common components, where $\boldsymbol{\Theta}$ is linked to the lower Cholesky factor of the trend innovation covariance matrix.

The co-integrated or common trend form has been extensively studied in the literature. Typically, in a given statistical application with an unknown data generating process, it is best to start with the open formulation in (14) and to entertain the possibility of common trends depending on the initial results. Where appropriate, the common trend form makes the model more parsimonious, which can lead to better parameter estimates and may improve its descriptive ability. Additionally, it entails an elegant structure in the multivariate filter, as in the illustration of the symmetric splitting of the gain function for the simulated example of the previous Section. The possibility of common slopes, that is the $m > 1$ case, has been less considered in empirical work, which has mostly focused on the I(1) case; there can be additional complexity in formal hypothesis testing for $m > 1$ compared to $m = 1$. At the same time, it is straightforward to evaluate statistical fit and diagnostic measures, such as those based on residual serial correlation, for the related versus common trends cases. Additionally, for multivariate filtering or signal detection applications, significant emphasis should be placed on considerations like extracted trend contour, implied gain functions, function-of-signal properties, and uncertainty in signal measurement.

## 4.2 Damped growth models

The above models give a straightforward and logical means to define multivariate generalizations of Butterworth filters and allow us to handle either related or common trends. At the same time, it is worthwhile to have models/filters capable of delivering smooth trends and signals with smooth growth rates, within a domain of I(1) forms. Here, we consider such models. A relatively sharp cutoff can still be attained in the low-pass filter; such behavior is made possible with a slope damping formula, applied recursively as shown in this sub-Section.

The $m-th$ order damped trend is given by

$$\begin{aligned} \boldsymbol{\mu}_{m,t} &= \boldsymbol{\mu}_{m,t-1} + \boldsymbol{\mu}_{m-1,t-1}, \\ \boldsymbol{\mu}_{i,t} &= diag(\boldsymbol{\phi})\boldsymbol{\mu}_{i,t-1} + \boldsymbol{\mu}_{i-1,t-1}, \quad i = 2, ..., m-1 \\ \boldsymbol{\mu}_{1,t} &= diag(\boldsymbol{\phi})\boldsymbol{\mu}_{1,t-1} + \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \sim NID(0, \boldsymbol{\Sigma}_\zeta), \end{aligned} \tag{16}$$

where $\boldsymbol{\mu}_{m-1,t}$ is an $N \times 1$ vector of slopes and $\boldsymbol{\Sigma}_\eta$ and $\boldsymbol{\Sigma}_{\boldsymbol{\xi}}$ are $N \times N$ positive semi-definite matrices. The coefficient vector $\boldsymbol{\phi}$ has all elements between zero and one. Often, all the $\phi_k' s$ have been set to unity, which gives the usual local linear trend. A more flexible trend results from allowing for values less than unity; such a damped slope was used in Trimbur (2010) to describe the tendency for a series to have its growth rate flatten out as the horizon extends. In this paper, we use the damped slope to allow for possible limits in the rates of sustained increases or decreases over time.

The extra flexibility in the stochastic trend specification permits them to remain based on I(1) specifications. This gives a modification to the Butterworth gain; the univariate form is

$$B_m^{lp}(\lambda) = \frac{1}{1 + \frac{1}{q_\zeta}(2 - 2\cos\lambda)(1 + \phi^2 - 2\phi\cos\lambda)^{m-1}}, \qquad 0 \leq \lambda \leq \pi,$$

To see where the gain equals one-half, we must solve

$$q_\zeta = (2 - 2\cos\lambda)(1 + \phi^2 - 2\phi\cos\lambda)^{m-1} \tag{17}$$

for $\lambda$ such that $0 \leq \lambda \leq \pi$. The solution $\lambda_{1/2}$ is in general a function of $m$, $q_\zeta$, and $\phi$. If we let $z = \cos\lambda$, then the equation becomes

$$(2-2z)(1+\phi^2-2\phi z)^{m-1} - q_\zeta = 0$$

and the solution for $z$ is given by the roots of an $m-$th degree of polynomial. These roots can be found numerically, though there is no analytical formula available. Values between -1 and 1 give $\lambda_{1/2}$ as the arc-cosine. Compared to the standard Butterworth, there is no compact expression for the gain in terms of $\lambda_{1/2}$.

However, the following proposition notes a useful approximation when $\phi$ is close to one, which will often be the case, as it gives rise to a relatively smooth trend.

**Proposition 1** *For $\phi \approx 1$, the cutoff frequency, at which the gain equals one-half, is approximately*

$$\lambda_{1/2} = 2\arcsin(2^{-1/2}[q_\zeta / r(m,\phi)]^{1/2m}) \tag{18}$$

*where $r(m,\phi) = 1 + (m-1)(-1+\phi)$*

In general, the low-pass filter does not completely annihilate the highest frequencies, though it may come very close. The minimal value occurs at $\lambda = \pi$ and equals

$$B_m^{lp}(\pi) = [1 + 4q_\zeta^{-1}(1+\phi^2-2\phi)^{m-1}]^{-1}$$

This may be seen by considering the difference

$$B_m^{lp}(\lambda) - B_m^{lp}(\pi) = \frac{4(1+\phi)^{2m-2} - (2-2\cos\lambda)(1+\phi^2-2\phi\cos\lambda)^{m-1}}{[q_\zeta + 4(1+\phi)^{2m-2}][1 + (2-2\cos\lambda)(1+\phi^2-2\phi\cos\lambda)^{m-1}/q_\zeta]}$$

This quantity is non-negative for all $\lambda$ and equals zero at $\lambda = \pi$. Therefore, the low-pass filter's gain function decays to a value that is very small but nonzero at $\lambda = \pi$; this minimal value decreases with $m$ and $q_\zeta$.

A canonical trend form can be obtained by replacing $\boldsymbol{\zeta}_t$ by $\boldsymbol{\zeta}_t + \boldsymbol{\zeta}_{t-1}$ in (16) and analogously $\boldsymbol{\mu}_{i-1,t-1}$ by $\boldsymbol{\mu}_{i-1,t} + \boldsymbol{\mu}_{i-1,t-1}$.

## 4.3 Adaptive matrix Filters for nonstationary time series

McElroy and Trimbur (2015) show that the Wiener-Kolmogorov formula generalizes to nonstationary vector processes, using pseudo-ACGF's in place of stationary ACGF's. Therefore, the optimal filter for extracting the trend from a doubly infinite series that follows (13) is

$$LP(L) = G_\mu(L)[G_\mu(L) + G_\varepsilon(L)]^{-1}. \tag{19}$$

Now for related trends that have uniform nonstationary operators across series, the pseudo-ACGF of the trend is

$$G_\mu(L) = f_\mu(L)\Sigma_{\boldsymbol{\zeta}}$$

for some scalar factor $f_\mu(L)$ ; the matrix ACGF of the irregular is $\Sigma_{\boldsymbol{\varepsilon}}$, so the filters take the form:

$$\begin{aligned} LP(L) &= \Sigma_\zeta f_\mu(L)[\Sigma_\zeta f_\mu(L) + \Sigma_\varepsilon]^{-1} \\ &= \Sigma_\zeta[\Sigma_\zeta + [f_\mu(L)]^{-1}\Sigma_\varepsilon]^{-1} \\ &= [(\Sigma_\zeta + [f_\mu(L)]^{-1}\Sigma_\varepsilon)\Sigma_\zeta^{-1}]^{-1}, \end{aligned}$$

which can be written as

$$LP(L) = [I_N + [f_\mu(L)]^{-1}\Sigma_\varepsilon\Sigma_\zeta^{-1}]^{-1}. \tag{20}$$

We will consider (13), where the signal of interest is $\mathbf{s}_t = \boldsymbol{\mu}_{m,t}$ for all $t$, with uniform differencing operator $\delta_{\mathbf{s}}(L) = (1-L)^m$ required to make the process white noise (In general, we consider the more flexible case where $\delta_{\mathbf{s}}(L) = (1-\phi L)^{m-1}(1-L)$, but here use the basic form with $\phi = 1$ for simplicity). For the m-th order trend, $f_\mu(L) = 1/[(1-L)(1-L^{-1})]^m$ is the pseudo-ACGF.

Now let $\boldsymbol{Q}$ be defined as $\Sigma_{\boldsymbol{\zeta}}\Sigma_{\boldsymbol{\varepsilon}}^{-1}$, which will be the multivariate generalization of the scalar signal-noise ratio $q$. Since $\Sigma_{\boldsymbol{\varepsilon}}$ is assumed invertible, $\boldsymbol{Q}$ always exists, even when $\Sigma_{\boldsymbol{\zeta}}$ is non-invertible. Then from (20), it follows that when $\Sigma_{\boldsymbol{\zeta}}$ is invertible,

$$LP(L) = \left(I_N + [(1-L)(1-L^{-1})]^m \boldsymbol{Q}^{-1}\right)^{-1}. \tag{21}$$

Formula (21) is a generalization of equation (7) for the univariate Butterworth filters with the scalar signal-noise ratio $q$ replaced by a matrix $\boldsymbol{Q}$. As such, (21) will not exist when $\Sigma_{\boldsymbol{\zeta}}$ is non-invertible because $\boldsymbol{Q}^{-1}$ is not well-defined. In the following, we assume that $\Sigma_{\boldsymbol{\zeta}}$ is invertible, which is the related trends case. The non-invertible case – with common trends – is treated in Appendix E and proceeds by re-expressing formula (21) in a way that does not involve $\Sigma_{\boldsymbol{\zeta}}^{-1}$ or $\boldsymbol{Q}^{-1}$. Here, we focus on the invertible related trends case, as the mathematics are simpler; the generalization to common trends is done in a similar manner. Note that in the special case that there is no cross-series correlation in either signal or noise, then $\Sigma_\zeta$ and $\Sigma_\varepsilon$ are diagonal matrices, and $\boldsymbol{Q}$ is a diagonal matrix consisting of signal-to-noise ratios for each of the component series.

In the frequency domain, the frequency response for the filter in (19) may be expressed through the spectra:

$$LP(\lambda) = F_\mu(\lambda)[F_\mu(\lambda) + F_\varepsilon(\lambda)]^{-1}, \tag{22}$$

where $F_\mu(\lambda) = G_\mu(e^{-i\lambda})$. Let $f_\mu(\lambda)$ denote the univariate trend's pseudo-spectrum for a unit variance disturbance. With homogenous operators the trend pseudo-spectrum is proportional to a constant density $f_\mu(\lambda) = 1/(2 - 2\cos\lambda)^m$, and we can write

$$\begin{aligned} LP(\lambda) &= \left(I_N + [2 - 2\cos\lambda]^m \boldsymbol{Q}^{-1}\right)^{-1} \\ &= \left(I_N + r(\lambda)\boldsymbol{Q}^{-1}\right)^{-1}, \end{aligned}$$

where $r(\lambda) = (2 - 2\cos\lambda)^m$.

We seek a formulation that has the same kind of expression as the original Butterworth sine filter – in terms of a cutoff frequency where the gain equals one-half – that was given in (1), but now using a vector of cutoff frequencies. Now let

$$\boldsymbol{Q}^{-1} = V\,\Lambda\,V^{-1}$$

be the eigen-decomposition of $\boldsymbol{Q}^{-1}$, with eigen-vectors given by the columns of the matrix $V$ and eigen-values given by the entries of the diagonal matrix $\Lambda$. The case of no cross-series correlation corresponds to $V = I_N$.

In general, the filter's gain is

$$\begin{aligned} LP(\lambda) &= \left(I_N + r(\lambda)V\,\Lambda\,V^{-1}\right)^{-1} \\ &= \left(V\,V^{-1} + r(\lambda)V\,\Lambda\;V^{-1}\right)^{-1}. \\ &= \left(V\,[V^{-1} + r(\lambda)\Lambda\;V^{-1}]\right)^{-1} \\ &= [V^{-1} + r(\lambda)\Lambda\;V^{-1}]^{-1}V^{-1} \\ &= [(I_N + r(\lambda)\Lambda)\;V^{-1}]^{-1}V^{-1}, \end{aligned}$$

which has the form

$$LP(\lambda) = V\,(I_N + r(\lambda)\Lambda)^{-1}\,V^{-1}. \tag{23}$$

One possible extension of the cutoff frequency concept to the multivariate case could heuristically be handled by setting $LP(\lambda)$ equal to $1/2$ times the identity matrix (denoted by $I_N$ for dimension $N$), in which case

$$V\left(I_N + r(\lambda_{1/2})\Lambda\right)^{-1} V^{-1} = (1/2)I_N$$

for some $\lambda_{1/2}$. Right-multiplying both sides by $V$ and then left-multiplying by $V^{-1}$ yields

$$\left(I_N + r(\lambda_{1/2})\Lambda\right)^{-1} = (1/2)I_N,$$

which implies

$$I_N = r(\lambda_{1/2})\Lambda.$$

The existence of a solution would mean that the eigenvalues in the diagonal of $\Lambda$ are all identical, so that $\boldsymbol{Q}^{-1}$ is a multiple of $I_N$. This only occurs in the very specific case that $\Sigma_{\boldsymbol{\zeta}}$ and $\Sigma_{\boldsymbol{\varepsilon}}$ are proportional.

Instead, we need a more general context. Suppose that a multivariate process $\{X_t\}$ is filtered to produce output $\{Y_t\}$, that is $Y_t = LP(L)\, X_t$. The spectral representation of the input process is

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t}\, dZ(\lambda)$$

where $Z(\lambda)$ is a vector-valued orthogonal increments process that indicates the differential (random) amplitude weighting the frequency component $e^{i\lambda t}$. The increment $dZ(\lambda)$ is in general complex-valued, though taking the integral from $-\pi$ to $\pi$, symmetrically about the origin, results in a cancellation of the imaginary part. The corresponding representation of the output process is

$$Y_t = \int_{-\pi}^{\pi} e^{i\lambda t}\, V\, (I_N + r(\lambda)\Lambda)^{-1}\, V^{-1}\, dZ(\lambda).$$

We can re-express these time series in terms of a rotated increments process, which is defined as $Z^{\star}(\lambda) = V^{-1}\, Z(\lambda)$:

$$X_t = \int_{-\pi}^{\pi} e^{i\lambda t}\, V\, dZ^{\star}(\lambda)$$
$$Y_t = \int_{-\pi}^{\pi} e^{i\lambda t}\, V\, (I_N + r(\lambda)\Lambda)^{-1}\, dZ^{\star}(\lambda).$$

The $j$th ($1 \leq j \leq N$) input and output series follow univariate orthogonal increment processes given by

$$\begin{aligned} X_{t,j} &= \mathbf{e}_j' \int_{-\pi}^{\pi} e^{i\lambda t} V dZ^{\star}(\lambda) \\ Y_{t,j} &= \mathbf{e}_j' \int_{-\pi}^{\pi} e^{i\lambda t} V\, (I_N + r(\lambda)\Lambda)^{-1}\, dZ^{\star}(\lambda) \end{aligned}$$

for $\mathbf{e}_j$ defined as a column vector with a one in the $j-$th position and zeroes elsewhere.

Another possible extension of the cutoff frequency idea could come from imposing that the $j$th output has half the amplitude or content of the $j$th input, or

$$\mathbf{e}_j'\, V\, (I_N + r(\lambda)\Lambda)^{-1} = \frac{1}{2}\, \mathbf{e}_j'\, V$$

for a total of $N$ equations. However, such a system will not in general have a solution. The above equation implies

$$(I_N + r(\lambda)\Lambda)^{-1} V' \mathbf{e}_j = \frac{1}{2}\, V' \mathbf{e}_j,$$

since $(I_N + r(\lambda)\Lambda)^{-1'} = (I_N + r(\lambda)\Lambda)^{-1}$. This requires that the vector $V'\, \mathbf{e}_j$ be an eigen-vector of $(I_N + r(\lambda)\Lambda)^{-1}$ with eigen-value $1/2$. However, since the matrix $(I_N + r(\lambda)\Lambda)^{-1}$ is diagonal, the only eigen-vectors in this case are unit vectors. Hence, the existence of solutions requires that $V'\,\mathbf{e}_j$ is a unit vector for all $j$, which solely holds for the special case of $V$ being a permutation matrix. This means that we cannot in general set the auto-gain equal to 1/2 for every multivariate filter applied to a vector process.

To generalize the cutoff frequency notion, we therefore impose a less stringent condition: for each $1 \leq j \leq N$, we find $\lambda_{1/2,j}$ such that the output has half the frequency content of the input when only the $j-$th process is incremented, that is

$$dZ^{\star}(\lambda) = dz^{\star}(\lambda)\mathbf{e}_j = \mathbf{e}_j dz^{\star}(\lambda),$$

where $dz^{\star}(\lambda)$ is a scalar orthogonal increments process. This corresponds to a hypothetical impulse whereby only series component $j$ passes through the filter. The equation for the amplitudes of frequency parts becomes

$$\mathbf{e}_j' V \left(I_N + r(\lambda_{1/2,j})\,\Lambda\right)^{-1} \mathbf{e}_j = \frac{1}{2}\mathbf{e}_j'\,V\,\mathbf{e}_j. \tag{24}$$

There are $N$ such equations. Further, $\lambda_{1/2,j}$ is allowed to vary over $j$; now collect the cutoff frequencies into a vector $\boldsymbol{\lambda}_{1/2} = (\lambda_{(1/2),1}, \lambda_{(1/2),2}, ..., \lambda_{(1/2),N})$. The $N$ equations reduce to scalar equations:

$$V_{j,j}\left(1 + r(\lambda_{1/2,j})\,\Lambda_{j,j}\right)^{-1} = \frac{1}{2}V_{j,j}\,, \quad j = 1, ..., N.$$

Each equation is solved by setting $r(\lambda_{1/2,j})\Lambda_{j,j} = 1$, or $r(\lambda_{1/2,j}) = 1/\Lambda_{jj}$, i.e., set the reciprocals of the eigen-values of $\boldsymbol{Q}^{-1}$ equal to $(2 - 2\,\cos(\lambda_j))^m$ and solve for $\lambda_j$.

The corresponding inverse eigenvalues can be seen as fundamental signal-noise ratios for the multivariate filter. That is, define

$$q_{\zeta,j} = 1/\Lambda_{jj}.$$

Then the solutions have the form

$$q_{\zeta,j} = (2 - 2\,\cos(\lambda_{1/2,j}))^m.$$

If we define

$$\lambda_{1/2,j} = 2\arcsin(2^{-1/2}{q_{\zeta,j}}^{(1/2m)}), \tag{25}$$

then

$$r(\lambda)\,\Lambda_{j,j} = [\sin(\lambda/2)/\sin(\lambda_{(1/2),j}/2)]^{2m}.$$

Let $\mathrm{S}(\lambda)$ be a diagonal matrix of dimension $N$ x $N$ with entry j, j given by $[\sin(\lambda/2)/\sin(\lambda_{(1/2),j}/2)]^{2m}$ for all $1 \leq j \leq N$. Since $\mathrm{S}(\lambda)_{j,j} = r(\lambda)\Lambda_{jj}$, the filter's gain is expressed as

$$LP(\lambda) = V\,(I_N + S(\lambda))^{-1}V^{-1}. \tag{26}$$

## 4.4 Bivariate Filters with correlated Signals

The parameter estimates for each series – the computed variances of driving disturbances – may improve when combined with shared parameters such as cross-series correlations. This more extensive information then becomes available in forming the signal extraction filters and enhances the filters' compatibility with series' dynamic properties. It also enables the researcher to more effectively pinpoint underlying signals in noisy series, as evidenced by a narrowing of confidence bands around smoothed trends. This enhancement is illustrated below in bivariate extraction of trends in oil consumption and imports time series. Application might feature a target series paired with another related series with stronger signal or a situation where two time series with comparable signal content are combined.

Now consider the bivariate case, labelling the series A and B. Therefore, the covariance matrices have the form:

$$\Sigma_\zeta = \begin{bmatrix} \sigma^2_{\zeta,A} & \sigma_{\zeta,AB} \\ \sigma_{\zeta,AB} & \sigma^2_{\zeta,B} \end{bmatrix}, \quad \Sigma_\varepsilon = \begin{bmatrix} \sigma^2_{\varepsilon,A} & \sigma_{\varepsilon,AB} \\ \sigma_{\varepsilon,AB} & \sigma^2_{\varepsilon,B} \end{bmatrix}.$$

Here, we can analytically compute the eigenvalues of the Signal-Noise Ratio matrix. First,

$$Q = \Sigma_{\boldsymbol{\zeta}}\Sigma_{\boldsymbol{\varepsilon}}^{-1} = (\sigma^2_{\varepsilon,A}\sigma^2_{\varepsilon,B} - \sigma^2_{\varepsilon,AB})^{-1} \begin{bmatrix} \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B} - \sigma_{\zeta,AB}\sigma_{\varepsilon,AB} & \sigma^2_{\varepsilon,A}\sigma_{\zeta,AB} - \sigma^2_{\zeta,A}\sigma_{\varepsilon,AB} \\ \sigma^2_{\varepsilon,B}\sigma_{\zeta,AB} - \sigma^2_{\zeta,B}\sigma_{\varepsilon,AB} & \sigma^2_{\zeta,B}\sigma^2_{\varepsilon,A} - \sigma_{\zeta,AB}\sigma_{\varepsilon,AB} \end{bmatrix}.$$

Denoting the elements of $Q$ by

$$Q = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

the equation for the eigenvalues becomes

$$\begin{vmatrix} \lambda_j - a & b \\ c & \lambda_j - d \end{vmatrix} = 0$$

or

$$\lambda_j^2 - trace(Q)\lambda_j + \det(Q) = 0$$

so that the solution is

$$\lambda_j = (1/2)trace(Q)\lambda_j \pm \sqrt{(1/4)[trace(Q)]^2 - \det(Q)} = 0.$$

Then in the case of contemporaneous trend relationships, the frequency response has all real elements, so the gain is the absolute value. Express the bivariate gain function as

$$G(\lambda) = \begin{bmatrix} G_A(\lambda) & G_{AB}(\lambda) \\ G_{BA}(\lambda) & G_B(\lambda) \end{bmatrix},$$

where $G_A$ is the gain applied to series A to estimate the component in A, and $G_{AB}(\lambda)$ is the cross-gain from B to A, that is, the gain applied to series B to estimate the component in series A. Let $\nu_{\zeta,AB} = \sigma_{\zeta,AB}/(\sigma_{\zeta,A}\sigma_{\zeta,B})$ equal the correlation between $\zeta_{A,t}$ and $\zeta_{B,t}$, and likewise for the irregular.

**Proposition 2** *The elements of the gain matrix for the bivariate Butterworth filter are given by*

$$\begin{aligned}
G_A(\lambda) &= h(\lambda)[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B} - \nu_{\zeta,AB}\nu_{\varepsilon,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}] \\
G_{AB}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,A}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B} + \nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,A}] \\
G_{BA}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,B}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B} + \nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,B}] \\
G_B(\lambda) &= h(\lambda)[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,B}\sigma^2_{\varepsilon,A} - \nu_{\zeta,AB}\nu_{\varepsilon,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}]
\end{aligned}$$

*where*

$$\begin{aligned}
h(\lambda) &= [f_\mu(\lambda)/R(\lambda)] = 1/[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + (1-\nu^2_{\varepsilon,AB})\sigma^2_{\varepsilon,A}\sigma^2_{\varepsilon,B}/f_\mu(\lambda) + (\sigma^2_{\varepsilon,A}\sigma^2_{\zeta,B} + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B}) \\
&\quad -2\nu_{\varepsilon,AB}\nu_{\zeta,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}\sigma_{\zeta,A}\sigma_{\zeta,B}].
\end{aligned}$$

This shows that the two cross-gains have the same parametric form in terms of $\lambda$ with different coefficients. Their ratio is

$$G_{AB}(\lambda)/G_{BA}(\lambda) = (-\sigma^2_{\zeta,A}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}+\nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,A})/(-\sigma^2_{\zeta,B}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}+\nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,B})$$

and so depends on all the variance and correlation parameters. Hence the contribution to signal A coming from filtering series B has a cross-gain that is a coefficient times a certain function of frequency, while the contribution to signal B from filtering series A is a different coefficient times the same quantity. For the special case $\sigma_{\varepsilon,AB} = 0$, the ratio is just $\sigma^2_{\varepsilon,A}/\sigma^2_{\varepsilon,B}$ and so depends only on the irregular variances independent of the trend disturbance's covariance matrix. Therefore, the cross-gains are equal when the irregulars are uncorrelated and $\sigma^2_{\varepsilon,A} = \sigma^2_{\varepsilon,B}$, regardless of the values of $\nu_{\zeta,AB}$, $\sigma_{\zeta,A}$, and $\sigma_{\zeta,B}$.

If, in addition $\nu_{\zeta,AB}$ is near one (but not exactly one), then away from the zero frequency, $G_A(\lambda)/G_B(\lambda) = q_{\zeta,A}/q_{\zeta,B}$; at very low frequences the gains de-couple and both $G_A(\lambda)$ and $G_B(\lambda)$

go to unity. In this case,

$$\begin{aligned}
G_A(\lambda) &= h(\lambda)[\sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B}] \\
G_{AB}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,A}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B} + \nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,A}] \\
G_{BA}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,B}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B} + \nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,B}] \\
G_B(\lambda) &= h(\lambda)[\sigma^2_{\zeta,B}\sigma^2_{\varepsilon,A}].
\end{aligned}$$

We can interpret the gain matrix as operating on the orthogonal increments, or Cramer representation. In general, the spectrum of the output is quadratic in $G(\lambda)$ and the cross-elements commingle with the full matrix spectrum of the input.

# 5 Extracting the trend-signal in petroleum consumption

In this application, interest centers on the trend in U.S. consumption of petroleum products, a key signal in several respects, given the direct share of overall consumption that petroleum demand represents and its indication of near-term economic activity. (Petroleum has a crucial role as an input and intermediate product in the industrial sector, and influences overall energy costs for consumer and industry.) There are also political dimensions to US oil consumption related to reliance on foreign oil and to conservation policies. However, the oil consumption time series is rather noisy, as indicated by the black line in Figure 7, making it difficult to ascertain the primary movements. The blue line represents an estimated trend from a statistical perspective, as explained below. Smoother trends may also be produced by higher order models, as in Figures 8 and 9; these have merit when the goal is to better track the long-run evolution. The red lines indicate uncertainty bounds, which can be expected to include the true value of the trend about 68% of the time. Clearly, any strategy for improving the trend estimate or its associated uncertainty is desirable.

Here, we use a second variable suggested by economic reasoning to inform the extracted trend in U.S. petroleum consumption and reduce the uncertainty connected with its measurement. In a bivariate analysis, we present gain patterns for different low-pass filter constructions, whereby both series are filtered with a certain weight function on the observations. There are various benefits to utilizing a secondary series in our trend assessment. First, when the trends in target and auxiliary series are cross-correlated, we can attain improved model performance and parameter estimates for both series. Second, richer information in the form of shared parameters becomes available, for instance the correlation between trend disturbances in the two series. Third, we can use their empirical relationships to improve the trend estimates in both series. When the supporting series also has high relative signal content – as in the following application – the formulas discussed earlier show how, in exact mathematical terms, to make the most efficient use of the series in enhancing the quality and precision of trend estimates in the target.

Specifically, the secondary series measures oil imports to the U.S. from the Organization of Petroleum Exporting Countries (OPEC). This imports series captures the marginal barrels of oil purchased as "last resort", following the more secure domestic and non-OPEC sources. Decisions on trade with the cartel, whose members include many politically unsettled nations, are often made on the margin and represent excess demand for oil. In this sense, imports from OPEC are indicative of existing economic conditions in petroleum markets and the strength of underlying demand, suggesting that they may help pin-point the trend in overall consumption. This intuitive reasoning is confirmed by the statistical analysis of the data that follows.

## 5.1 Data

The time series data are monthly Total U.S. Consumption of Petroleum Products, Industrial Sector, and U.S. Petroleum Imports from OPEC countries, measured in Thousands of Barrels per Day. The data are taken from the Department of Energy, and the sample period is January 1973 to February 2018. All models and filters are applied to the logarithms of each series.

The original data are clearly seasonal. For example, the seasonality shows up strongly in the residual serial correlations for models fitted with only trend and noise. However, the data are available to the public only as unadjusted data; there is no official seasonal adjustment conducted. Therefore, prior to our trend-noise analysis, we seasonally adjusted each series via a model-based approach, as detailed in an Appendix. All calculations were carried out using programs written in the Ox language of Doornik (2013) designed for statistical and econometric applications. The programs made use of the SSfPack package of state space and Kalman filter functions that is described in Koopman et. al (2008).

## 5.2 Univariate Results with Standard Forms

For the standard univariate case, model (13) was fitted along with (16) using a minimum value for $\phi$ of 0.95. Parameter estimates are reported in Table 1 for orders $m$ from 1 to 5. The table indicates the core slope disturbance variance $\sigma_{\zeta}^2$ (the level disturbance when $m = 1$) along with the irregular variance $\sigma_{\epsilon}^2$. As the order increases, $\sigma_{\zeta}^2$ decreases regularly while $\sigma_{\epsilon}^2$ rises. The gains in the irregular variance with $m$ mean that additional noise is increasingly removed from the trend, which therefore becomes progressively smoother. The incremental growth in $\sigma_{\epsilon}^2$ tapers off as the order advances, and once the trend order reaches around 4 or 5, further increases in $m$ produce little change in the trend's shape. The signal-noise ratio defined as $q_\zeta = \sigma_\zeta^2/\sigma_\epsilon^2$ is a scale invariant measure that, for a given order, represents the relative variation in signal process versus irregular movements. Within each component structure such a ratio can be used to compare dynamics across series that have different scales.

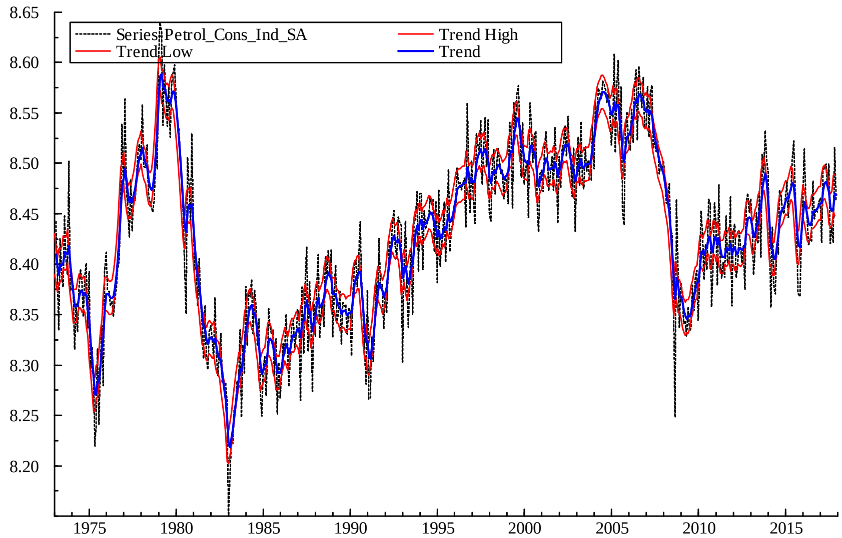


Figure 7: Estimated trend in Petroleum Consumption for univariate first-order standard model.

The estimates for $q_\zeta$ reported in Table 1 decline steadily as $m$ ranges from 1 to 5, with substantial reductions for each increment in order. The value of $q_\zeta$ for $m = 5$ is about 2 x $10^{-9}$ times the value for the first order case. Results are shown only up to $m = 5$ partly because in the bivariate case, the numerical optimization of the Log-Likelihood function is infeasible for orders above 5 (due to very small values of $\sigma_\zeta^2$ combined with the high dimensionality of the state space) and partly because such higher orders yield very little difference in the visible trend compared to the fifth order model. Therefore, to be comparable to the bivariate tables, the univariate results are shown up to fifth order but no larger. In Table 1, $\overline{\beta}$ indicates the measured constant drift for the random walk plus drift trend or, for all $m$ larger than one, the unconditional mean of the slope process underpinning the trend (which itself has the form of an order $m - 1$ process). The damping coefficient $\phi$ is shown in the last column of the table; for all orders above one the coefficient is well-defined, and the estimates go to the minimum permissible value in each case.

The application of the signal extraction formulas, given the estimated parameters, yields the consumption trend shown in Figure 7 for first order. The trend meanders throughout the sample, its general level evolving slowly over the sample period; it also undergoes frequent changes in direction and small adjustments on a quarterly basis. Such an adaptive level might be intuitive to the extent that underlying consumption is affected by various factors that are constantly changing and that

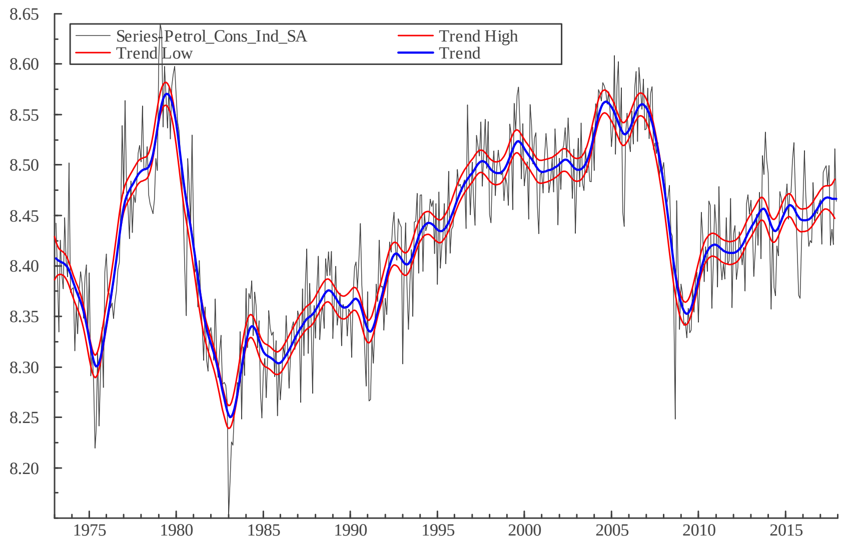


Figure 8: Estimated trend in Petroleum Consumption for univariate second-order standard model.

Table 1: Maximum likelihood estimates of parameters for Series: 'Petrol_Cons_Ind_SA'

| Trend | ZetaVar | EpsVar | Q_Zeta | BetaMean | Phi |
|---|---|---|---|---|---|
| Standard (Order 1) | 0.0003322 | 0.0009667 | 0.3436 | 0.000109 | N/A |
| Standard (Order 2) | 6.33e-006 | 0.001361 | 0.00465 | 0.0001502 | 0.95 |
| Standard (Order 3) | 5.588e-008 | 0.001531 | 3.65e-005 | 0.0006025 | 0.95 |
| Standard (Order 4) | 2.355e-010 | 0.001671 | 1.409e-007 | -0.0002134 | 0.95 |
| Standard (Order 5) | 1.079e-012 | 0.00175 | 6.167e-010 | -0.001894 | 0.95 |

Note: The Model Type for Observations 'Standard (Order m) + Irregular' indicates an Unobserved Components model with Standard m-th order Stochastic Trend plus a White Noise Irregular.

'Canonical' indicates a Canonical Stochastic Trend.

may have permanent effects. Likewise, the resulting trends are shown in Figure 8 for second order and in Figure 9 for fifth order. The increased smoothness of these higher order cases better clarifies the primary changes in level over the sample period. Such trends may be appealing, as there seems to be less of a response to temporary patterns. The estimated signal for $m = 5$ provides an even smoother trajectory than the second order cases and makes it easy to see the most important changes in the level of consumption over time-scales of a decade or more. Thus, in Figure 9, the declines in the mid-70s and early 80s show up very clearly, as do the long periods of increase during the mid- to late-80s, the 90s, and around the mid-2000s.

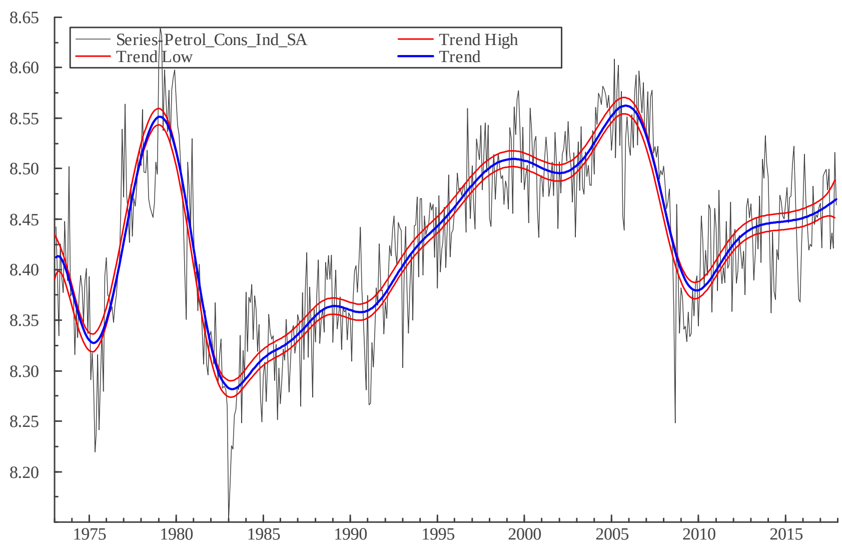


Figure 9: Estimated trend in Petroleum Consumption for univariate fifth order standard model.

The corresponding gain functions are shown in figure 10; they show directly how the higher frequency parts are more effectively cut out when the order is increased. The gain for the $m = 1$ case decays far more gradually than that of the fifth order case. From the standpoint of evaluating low-pass filters according to their ability to remove the high end of the spectrum, the higher orders may be seen as desirable. However, from another perspective, with stochastic components that have spectral overlap, flexibility in the gain such as that of the first order model arises quite naturally and may be a positive attribute; this produces a level that adapts more quickly to change. Depending on the application, it might become beneficial to employ a broader interpretation of low pass filtering that accomodates gain functions with slow descent. The higher order cases that produce shapes more like the standard low-pass filters allow one to focus better on the most essential evolution in the level over the entire sample period. In parallel with the gain's ability for noise removal, the estimated trend curves start to converge as the order reaches around 5 or so. Higher orders such as $m = 8$ (which was feasible to estimate in the univariate case) yield very little visible differences

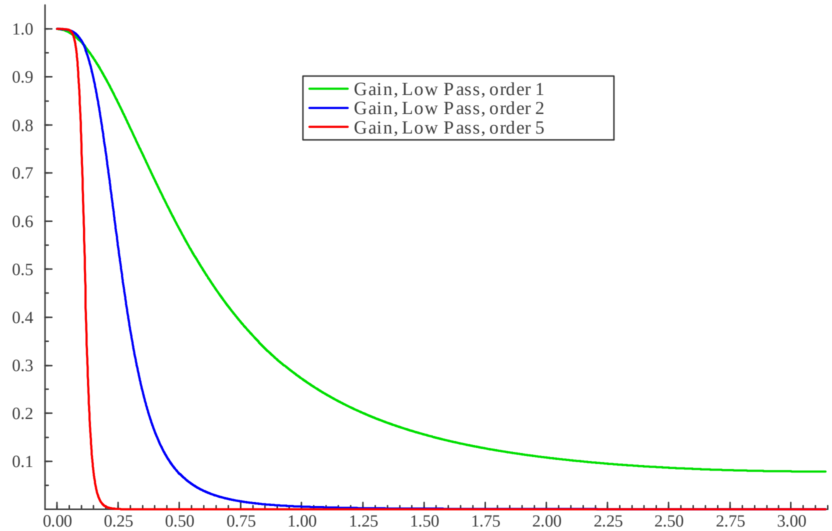


Figure 10: Estimated gain function for extracting the trend in Petroleum Consumption for univariate standard models of first, second, and fifth order.

in the trend from a fifth order specification.

Table 2: Estimated Trend and RMSE values for Series: 'Petrol_Cons_Ind_SA'

| Trend | Trend-Mid | RMSE-Mid | RMSE-Pct | Trend-End | RMSE-End | RMSE-Pct |
|---|---|---|---|---|---|---|
| Standard (Order 1) | 8.447 | 0.0165 | 0.1952% | 8.4686 | 0.0206 | 0.243% |
| Standard (Order 2) | 8.443 | 0.0113 | 0.1336% | 8.4665 | 0.0195 | 0.230% |
| Standard (Order 3) | 8.435 | 0.0096 | 0.1139% | 8.4699 | 0.0193 | 0.227% |
| Standard (Order 4) | 8.434 | 0.0086 | 0.1018% | 8.4705 | 0.0187 | 0.221% |
| Standard (Order 5) | 8.435 | 0.0080 | 0.0951% | 8.4698 | 0.0186 | 0.219% |

Note: Trend estimates around the sample mid-point and sample end-point along with Root-Mean-Square-Errors (RMSEs) are indicated in the table above.

Figures 7, 8, and 9 also show confidence bands around the trend estimates, constructed by adding or subtracting the Root-Mean Square Error (RMSE), conditional on the parameter values being set to the Maximum Likelihood Estimates. (In making comparisons across different orders, it should be borne in mind that the size of the bands conditions on the particular order being the correct one in each case.) It appears that, away from end-points, confidence intervals narrow as order increases. For larger $m$ the uncertainty tends to rise as the estimation point approaches the end of the series.

This pattern is reflected in Table 2, which reports trend estimate values and their RMSE, in addition to the percentage of the estimate that the RMSE represents, around the mid-point of the sample (Jan 1994) and at the end-point (Feb 2018). The mid-series confidence bands decrease in size with increasing order, with the biggest reduction occurring when the order goes from one to two; for $m = 5$ the RMSE is roughly one-half that for $m = 1$. As can be seen in Figures 7 to 9, the uncertainty in trend estimate grows in moving from mid-series to near the end-point; according to Table 2, for $m = 3$ the RMSE roughly doubles from middle to end of sample.

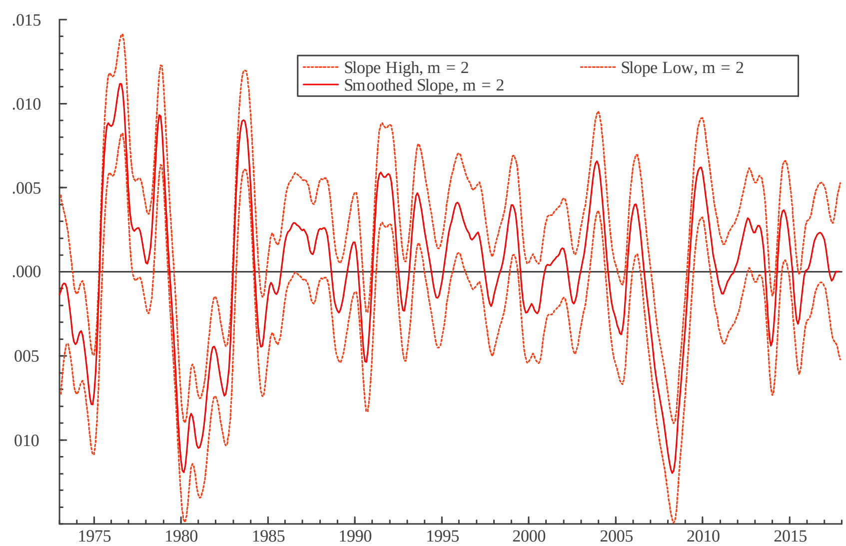


Figure 11: Estimated slope of signal in Petroleum Consumption for univariate second-order standard model, along with 67% confidence bands.

Model fit and diagnostic statistics are shown in an Appendix. $R_D^2$ is the coefficient of determination relative to first differences, which is analogous to the usual regression $R^2$ that applies to stationary data. The quantity $\widehat{\sigma}$ is the equation standard error. $Q(P)$ is the Box-Ljung statistic composed of $P$ residual autocorrelations. Three different values of $P$ are considered for these statistics; $Q(P)$ should be compared with a $\chi^2$ distribution with $P-2$ degrees of freedom. The Akaike Information Criterion (AIC) is $AIC = -2\log\hat{L} + 2k$, where $\log\hat{L}$ is the maximized log-likelihood for each specification, with $k$ the number of model parameters. Likewise, the Schwarz Information Criterion (SIC) is $SIC = -2\log\hat{L} + (\log T)k$, where $T$ is the sample size. The $Q(P)$ values rise

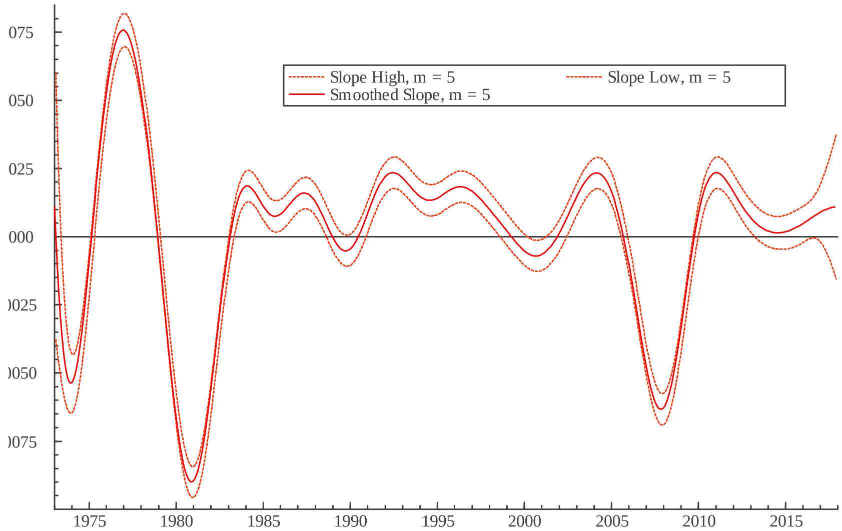


Figure 12: Estimated slope of signal in Petroleum Consumption for univariate fifth-order standard model, along with 67% confidence bands.

appreciably with increasing order, while the coefficient of fit $R_D^2$ fall. $Q(36)$ increases modestly from 114 to around 124 in moving from first to second order, and then jumps to about 240 for $m = 5$. If one has a desire for smoothness, then the second order model may represent a suitable compromise between pure model fit and regularity of trend. Further discussion is contained in an Appendix.

The slope, or systematic rate of change in the signal, undergoes time variation for $m = 2$ and higher; the estimated slope series can provide useful information about the contour of a series' historical evolution and its near end-point direction. The second order case is displayed in Figure 11, while the fifth order slope is plotted in Figure 12; the 67% confidence bands are indicated as well. Similarly to the smoother path of the trend itself, the slope for $m = 5$ appears far more fluid and stable than for $m = 2$. Superposition of the two slopes reveals that the fifth order slope tends to pass through the swings in the second order one. Further, the confidence bands in figure 12 are substantially reduced in size relative to those in Figure 11 (again these depend on the model being correct.)

Table 3: Maximum likelihood estimates of parameters for Series: 'Oil Imports from OPEC'

| Trend | ZetaVar | EpsVar | Q_Zeta | BetaMean | Phi |
|---|---|---|---|---|---|
| Standard (Order 1) | 0.005363 | 0.000566 | 9.475 | 0.0001263 | N/A |
| Standard (Order 2) | 0.001029 | 0.002862 | 0.3596 | 0.0007004 | 0.95 |
| Standard (Order 3) | 2.989e-006 | 0.006475 | 0.0004617 | -0.00284 | 0.95 |
| Standard (Order 4) | 9.215e-008 | 0.006846 | 1.346e-005 | -0.007782 | 0.95 |
| Standard (Order 5) | 2.982e-009 | 0.007089 | 4.207e-007 | 0.01029 | 0.95 |

See notes to table 1.

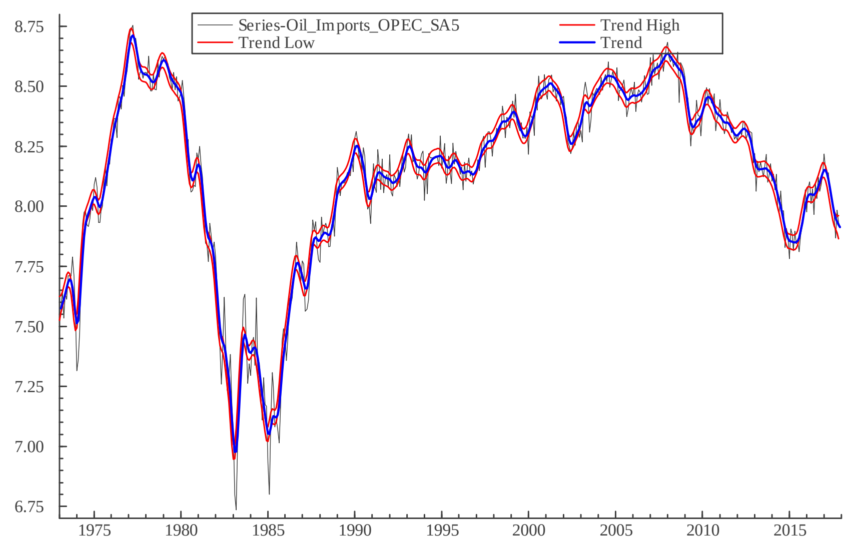


Figure 13: Estimated trend in Crude Oil Imports from OPEC for univariate second-order standard model.

As shown in Table 3, the estimated $q$ values for the time series of OPEC imports are substantially higher than those for consumption; this comparison of ratios gives a statistical description of the greater "signal content" in imports and indicates possibilities for enhancing the targeted signal in consumption. As shown in Figure 13, the series is somewhat noisy but mainly undergoes highly persistent movements. Most of its variation appears to be accounted for by the stochastic trend, shown in blue. The temporary noise that represents fluctuations around the trend is easily seen to be of smaller relative range than for the consumption series.

## 5.3 Univariate Results with Canonical Forms

We also estimated the canonical models discussed earlier. Tables of results are contained in Appendix A. The trend disturbance variance is substantially smaller for the canonical compared to the standard case, owing to reinforcement by the MA part of the process. The irregular variance is larger for the canonical version for orders 1; for higher orders the standard model's irregular has about the same variance. The damping coefficient again goes to the boundary value of 0.95 for orders 2 and higher. As shown in Figures B.3 and B.4 in the Appendix, which focus on sub-samples for visibility, the canonical trend is somewhat smoother as it mitigates some of shorter-term gyrations in the standard trend. For higher orders, the differences in the two trends are negligible. The gains for the first order case are displayed in Figure B.5; the augmented smoothness in the canonical trend is connected with the greater removal of high frequencies apparent in its gain function.

## 5.4 Bivariate Modelling

For the bivariate model, the parameter estimates for oil consumption are shown in Table 4; the variance of the trend disturbance increases noticeably compared to the univariate model. We may interpret this result as indicating that the bivariate model with correlated trend disturbances helps to "draw out" additional systematic variation in consumption; a more variable trend comes from linking it to the highly flexible OPEC imports trend. The estimated cross-correlations are reported in Table 5. The trend's correlation parameter rises sharply from first to second order to about 0.94, and then ranges from 0.84 to 0.98 for higher orders. The tight link between trends is combined with more or less unrelated movements in the irregular. In comparing Table A.1, which contains diagnostics for the univariate standard case, with Table A.2 that pertains to the bivariate extension, we see a substantial improvement in the Q-statistics and a moderate increase in the coefficient of determination relative to first differences.

The RMSEs of the trend estimates are presented in Table 5; these values are mostly lower than the ones reported in Table 2. The differences are not large, because the trend is assessed as being more variable in the bivariate case. In terms of trend estimation, it is important to note the alteration in individual variance parameter from univariate to bivariate, and a coherent comparison involving the bivariate trend's degree of uncertainty requires taking this change into account. Other factors being the same, a larger trend shock variance has a positive impact on MSE in moving to the bivariate case, but such an increase is not necessarily undesirable; it is a by-product of a more

Table 4: Maximum likelihood estimates of parameters, Series A: 'Petrol_Cons_Ind_SA'

| Trend | ZetaVar | EpsVar | Q_Zeta | BetaMean | Phi |
|---|---|---|---|---|---|
| Standard (Order 1) | 0.0003384 | 0.000976 | 0.3468 | 0.0001347 | N/A |
| Standard (Order 2) | 3.695e-005 | 0.001277 | 0.02893 | 0.0003773 | 0.95 |
| Standard (Order 3) | 1.031e-007 | 0.00152 | 6.782e-005 | 0.0004989 | 0.95 |
| Standard (Order 4) | 1.918e-009 | 0.001613 | 1.189e-006 | -0.001045 | 0.95 |
| Standard (Order 5) | 2.784e-011 | 0.001692 | 1.645e-008 | -0.0007123 | 0.95 |

See notes to table 1.

Table 5: Shared STSM Model parameters

Bivariate Models, Series A: Petrol_Cons_Ind_SA. Series B: Oil_Imports_OPEC_SA_LagOne.

| Model Type for Observations | ZetaCorr | EpsCorr |
|---|---|---|
| Standard (Order 1) + Irregular | 0.5428 | -0.5055 |
| Standard (Order 2) + Irregular | 0.9472 | -0.1447 |
| Standard (Order 3) + Irregular | 0.8425 | 0.1059 |
| Standard (Order 4) + Irregular | 0.9561 | 0.114 |
| Standard (Order 5) + Irregular | 0.9832 | 0.1326 |

Note: ZetaCorr represents the correlation between trend disturbances, while EpsCorr represents the irregular correlation.

flexible trend that improves the statistical fit of the model.

To assess the pure impact of the bivariate linkages, we can control for the individual variance parameter and isolate the adjustment in MSE coming from the non-zero correlations across series. This is achieved by considering a univariate specification with variances set to the bivariate model estimates. Such a specification yields the estimated trend values and uncertainty measures shown in Table 7. It can be seen that there are significant reductions in uncertainty coming from the bivariate relations; for instance, the MSE decreases by over 35% for the second order case.

Table 6: Estimated Trend and RMSE values for Series A: 'Petrol_Cons_Ind_SA'

| Trend | Trend-Mid | RMSE-Mid | RMSE-Pct | Trend-End | RMSE-End | RMSE-Pct |
|---|---|---|---|---|---|---|
| Standard (Order 1) | 8.445 | 0.0155 | 0.1835% | 8.4654 | 0.0196 | 0.231% |
| Standard (Order 2) | 8.443 | 0.0111 | 0.1320% | 8.4620 | 0.0188 | 0.222% |
| Standard (Order 3) | 8.433 | 0.0093 | 0.1108% | 8.4568 | 0.0186 | 0.220% |
| Standard (Order 4) | 8.433 | 0.0085 | 0.1010% | 8.4523 | 0.0185 | 0.219% |
| Standard (Order 5) | 8.434 | 0.0080 | 0.0952% | 8.4525 | 0.0186 | 0.220% |

The bivariate gain function away from the end-points is shown in Figure 14. The trend in consumption is estimated by applying a relatively sharp low-pass filter to its own series, in addition to a cross-filter focussed on an intermediate span of frequencies being applied to the imports series. As noted above, the two cross-filters have gains that are proportional to each other, and since the irregular correlation is relatively small, the coefficient of proportionality is about equal to the irregular variance ratio. The auto-gain for OPEC imports admits considerably more of the medium-frequency region of the spectrum than that of the consumption series. The auto-gain for consumption is relatively low since the bivariate filter is relying on the ancillary series for an important share of the input.

In the above illustration, just one additional series is used to attain a substantial decrease in uncertainty. Its key attributes are that its trend is highly correlated with the target's, its irregular is roughly uncorrelated, and it has a high signal-noise ratio. Here, we have used the simplest possible model that keeps the I(1) feature across specifications; other features such as ARMA sampling errors, stochastic cycles, or other components could be incorporated as well.

Table 7: Estimated Trend and RMSE values for Series: 'Petrol_Cons_Ind_SA'

| Trend | Trend-Mid | RMSE-Mid | RMSE-Pct | Trend-End | RMSE-End | RMSE-Pct |
|---|---|---|---|---|---|---|
| Standard (Order 1) | 8.447 | 0.0166 | 0.1966% | 8.4687 | 0.0207 | 0.245% |
| Standard (Order 2) | 8.448 | 0.0139 | 0.1640% | 8.4648 | 0.0232 | 0.274% |
| Standard (Order 3) | 8.437 | 0.0101 | 0.1198% | 8.4690 | 0.0203 | 0.240% |
| Standard (Order 4) | 8.434 | 0.0097 | 0.1155% | 8.4694 | 0.0217 | 0.256% |
| Standard (Order 5) | 8.432 | 0.0095 | 0.1124% | 8.4696 | 0.0228 | 0.269% |

Note: See notes to table 2. Here, trend estimates and RMSEs are reported for parameter values set to those in table 4.

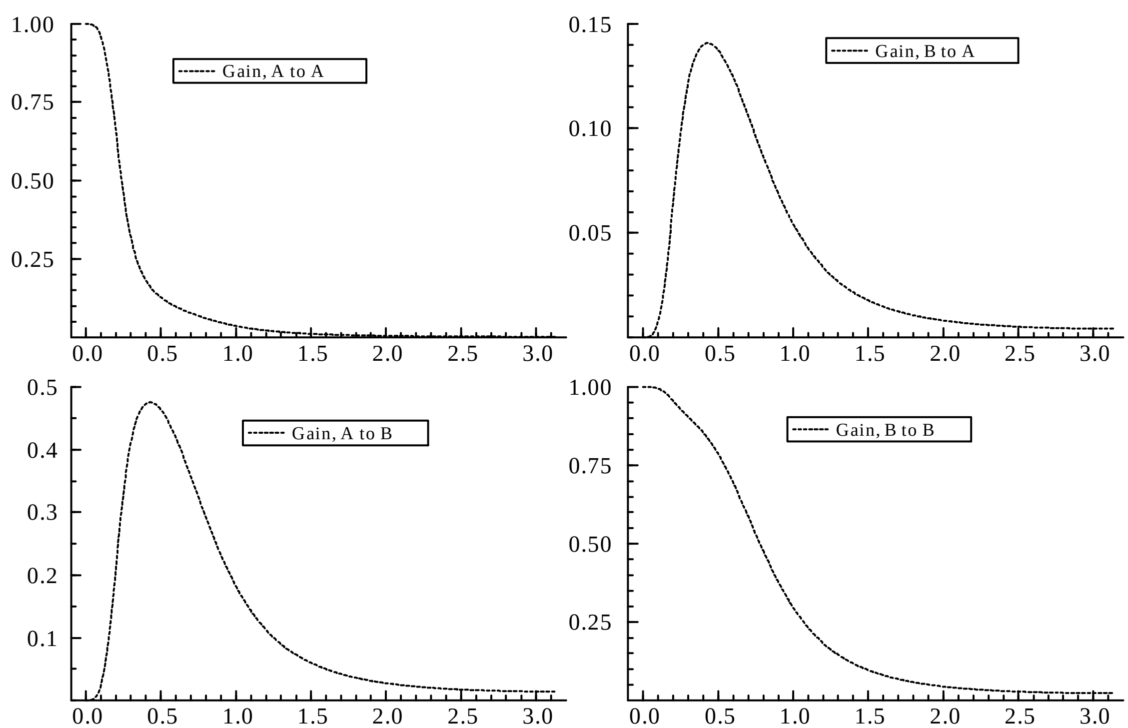


Figure 14: Bivariate Gain function for dataset on Petroleum Consumption and OPEC imports.

# 6 Conclusions

A vast range of examples has appeared in scientific literature where researchers have used methods for estimating regular components in series subject to noisy fluctuations. Butterworth filters are one of the most commonly used classes for this purpose. The key limitations that this paper aims to remedy are their restriction to single-series cases and their application as ad hoc filters. Scientists often work with multivariate datasets consisting of observations on related variables; making the best use of such data can improve accuracy and reduce error in detecting signals. Also, since key interpretations and conclusions about research findings can depend importantly on how the data's regularities are measured, it is essential to use a consistent measurement approach supported by evidence.

This paper provides multivariate generalizations of the class of Butterworth filters by way of their rigorous basis in statistical modelling. Previous work has used Butterworth and related filters to eliminate noise in diverse applications, yet only univariate cases have been considered, with the filters designed without reference to the data's properties. Even in multivariate applications, Butterworth filters have been used only in univariate fashion.

We develop formulas for the multivariate low-pass filters by proceeding from the model-based representation of existing Butterworth filters, originally revealed in the economic statistics literature. In the frequency domain, the multivariate filters have gain functions with similarly flexible shapes in terms of cutoff location and sharpness and include cross-gain elements that show exactly how a measurement of a target signal is informed by other, related series. The multivariate filters can yield more refined signals by combining the information in related series. Through estimation or calibration of model parameters, the filter design respects the empirical interrelationships across series, in particular the co-movements in systematic versus non-systematic fluctuations.

These features are illustrated by the use of OPEC oil imports to help gauge the trend in oil consumption. There are numerous possible applications in future work that may include time series in fields such as oceanography, climatology, and astronomy – in situations where standard univariate filters are being used for related series, the methodology used may be expanded to the multivariate filters. In applications with greater numbers of series, we might achieve even greater gains in signal accuracy when using series with the three key properties of correlated trend, uncorrelated irregular, and signalling ability.

# References

[1] Adeyemi, O. I., Broadstock, D. C., Chitnis, M., Hunt, L. C., & Judge, G. (2010). Asymmetric price responses and the underlying energy demand trend: Are they substitutes or complements?

Evidence from modelling OECD aggregate energy demand. *Energy Economics* **32(5)** 1157-1164.

[2] Agnolucci, P. (2010). Stochastic trends and technical change: the case of energy consumption in the British industrial and domestic sectors. *The Energy Journal* **1** 111-135.

[3] Avdeeva, D. K., Kazakov, V. Y. E., Natalinova, N. M., Ivanov, M. L., Yuzhakova, M. A., & Turushev, N. V. (2014). The simulation results of the high-pass and low-pass filter effect on the quality of micropotential recordings on the electrocardiogram. *Biology and Medicine* **6(1)** 1-10.

[4] Bell, W. (1984) Signal extraction for nonstationary time series. *The Annals of Statistics* **12** 646–664.

[5] Bell, W. (2004) On RegComponent time series models and their applications. In A.C. Harvey, S.J. Koopman, and N. Shephard (Eds.), *State Space and Unobserved Component Models: Theory and Applications*. Cambridge, UK: Cambridge University Press.

[6] Broadstock, D. C., Collins, A., & Hunt, L. C. (2011). Transportation oil demand, consumer preferences and asymmetric prices. *Journal of Economic Studies* **38(5)** 528-536.

[7] Butterworth, S. (1930). On the theory of filter amplifiers. *Wireless Engineer* **7(6)** 536-541.

[8] Coumou, D., Robinson, A., & Rahmstorf, S. (2013). Global increase in record-breaking monthly-mean temperatures. *Climatic Change* **118(3-4)** 771-782.

[9] Dermoune, A., Djehiche, B., & Rahmania, N. (2009). Multivariate extension of the Hodrick-Prescott filter-optimality and characterization. *Studies in Nonlinear Dynamics & Econometrics* **13(3)** 1-33.

[10] Dimitropoulos, J., Hunt, L. C., & Judge, G. (2005). Estimating underlying energy demand trends using UK annual data. *Applied Economics Letters* **12(4)** 239-244.

[11] Doornik, J. A. (2013), *Object-Oriented Matrix Programming Language Using Ox,* London: Timberlake Consultants Ltd.

[12] Durbin, J. and Koopman, S. (1997) *Time Series Analysis by State Space Methods*, volume 24 of *Oxford Statistical Science Series*. Oxford: Oxford University Press.

[13] Gallant, A. J., Karoly, D. J., & Gleason, K. L. (2014). Consistent trends in a modified climate extremes index in the United States, Europe, and Australia. *Journal of Climate* **27(4)** 1379-1394.

[14] Gao, H., Zhang, B. B., & Zhang, B. (2012). Stepwise filter correlation method and evidence of superposed variability components in gamma-ray burst prompt emission light curves. *The Astrophysical Journal* **748(2)** 1-14.

[15] Gehring, D., Wissler, S., Mornieux, G., & Gollhofer, A. (2013). How to sprain your ankle–a biomechanical case report of an inversion trauma. *Journal of Biomechanics* **46(1)** 175-178.

[16] Gomez V. (2001). The use of Butterworth filters for trend and cycle estimation in economic time series. *Journal of Business and Economic Statistics* **19** 365-73.

[17] Harvey, A. C. (1989), *Forecasting, structural time series models and the Kalman filter*, Cambridge: Cambridge University Press.

[18] Harvey, A. C., and Koopman, S. J. (1997). Multivariate structural time series models. In *System dynamics in economic and financial models* (eds C. Heij et al.), Chichester: Wiley and Sons.

[19] Harvey, A. C., and Trimbur T. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics* **85** 244–55.

[20] Hillmer, S. C., & Tiao, G. C. (1982). An ARIMA-model-based approach to seasonal adjustment. *Journal of the American Statistical Association* **77(377)** 63-70.

[21] Hunt, L. C., Judge, G., & Ninomiya, Y. (2003). Underlying trends and seasonality in UK energy demand: a sectoral analysis. *Energy Economics* **25(1)** 93-118.

[22] Koopman, S. J., Shephard N., and Doornik J. (2008), *Statistical Algorithms for Models in State Space: SsfPack 3.0*, London: Timberlake Consultants Ltd.

[23] McElroy, T. and Trimbur, T. (2015). Signal Extraction for Non-Stationary Multivariate Time Series with Illustrations for Trend Inflation. *Journal of Time Series Analysis* **36** 209-227

[24] McElroy, T. (2008) Matrix formulas for nonstationary ARIMA signal extraction. *Econometric Theory* **24** 1–22.

[25] McElroy, T. (2017). Multivariate seasonal adjustment, economic identities, and seasonal taxonomy. *Journal of Business & Economic Statistics* **35(4)** 611-625.

[26] Nyblom, J., & Harvey, A. (2000). Tests of common stochastic trends. *Econometric Theory* **16(2)** 176-199.

[27] Poloni, F. and Sbrana, G., (2017). Multivariate Trend–Cycle Extraction With The Hodrick–Prescott Filter. *Macroeconomic Dynamics* **21(6)** 1336-1360.

[28] Stock, J. H., & Watson, M. W. (1988). Testing for common trends. *Journal of the American Statistical Association* **83(404)** 1097-1107.

[29] Tauzin, B., Debayle, E., & Wittlinger, G. (2010). Seismic evidence for a global low-velocity layer within the Earth's upper mantle. *Nature Geoscience* **3(10)** 718.

[30] Tiao, G. C., & Hillmer, S. C. (1978). Some consideration of decomposition of a time series. *Biometrika* **65(3)** 497-502.

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

[32] Trimbur, T. M. (2010). Stochastic level shifts and outliers and the dynamics of oil price movements. *International Journal of Forecasting* **26(1)** 162-179.

[33] Vahid, F., & Engle, R. F. (1993). Common trends and common cycles. *Journal of Applied Econometrics* **8**(**4**) 341-360.

[34] Visser, H., Dangendorf, S., & Petersen, A. C. (2015). A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate". *Journal of Geophysical Research: Oceans* **120(6)** 3873-3895.

[35] Visser, H., Dangendorf, S., Van Vuuren, D. P., Bregman, B., & Petersen, A. C. (2018). Signal detection in global mean temperatures after "Paris": an uncertainty and sensitivity analysis. *Climate of the Past* **14(2)** 139.

[36] Whittle, P. (1963) *Prediction and Regulation*. London: English Universities Press.

[37] Wiener, N. (1949) *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. New York: Wiley.

# Appendix A Tables of Estimation Results

The values of $Q(P)$ are beyond the critical values for 1 or 5 percent levels of significance for the consumption series; this stems mainly from the third order autocorrelation being non-zero. We experimented with extensions of the model, such as versions with a negative-coefficient AR(1) that absorbs very high frequency movements, that offer some reduction in residual serial correlation. It appears that more expansive forms of models, such as those with stochastic cycle or ARMA process components, could be used to produce more satisfying values of Q(P) statistics; however, such explorations would deviate from the main goals of the current paper. Here we focus on the standard trend plus irregular models, as these are the natural starting point and the basis for connections with Butterworth filters that pave the way for multivariate generalization, which has not yet been provided in the literature to date.

Table A1: Diagnostics for Series: 'Petrol_Cons_Ind_SA'

| Trend | Q(12) | Q(24) | Q(36) | Eq SE | AIC | SIC | R^2_d |
|---|---|---|---|---|---|---|---|
| Standard (Order 1) | 45.47 | 96.31 | 127.9 | 0.0415 | -1885.37 | -1885.37 | 0.2451 |
| Standard (Order 2) | 108 | 163.4 | 198 | 0.0435 | -1833.53 | -1833.53 | 0.1727 |
| Standard (Order 3) | 132.5 | 193.9 | 222.1 | 0.045 | -1794.72 | -1794.72 | 0.1146 |
| Standard (Order 4) | 149.3 | 245.9 | 267.4 | 0.046 | -1769.77 | -1769.77 | 0.07392 |
| Standard (Order 5) | 164.1 | 284.1 | 304 | 0.0467 | -1753.53 | -1753.53 | 0.04556 |

Note: AIC = -2*LogL_Max +2*k, where k is number of parameters. SIC = -2*LogL_Max +2*log(T)*k for series length T. R^2_x is the coeff. of determination relative to simple benchmark, a RW (with fixed seasonal dummies for seasonal data. Specifically, R^2_x = 1-PEV(model)/PEV(benchmark), where PEV is the Prediction Error Variance (in KF steady state).

For the OPEC imports series, the fit appears best for the first order case, and the diagnostics worsens for higher orders. The Q-statistics suggest serial correlation in the residuals, which could possibly be addressed by using serially correlated components in addition to the trend and noise components. However, our focus here is on the signal extraction and filtering problem in the mlutivariate setting, and on making connections with Butterworth filters. The main issues are clearest for the simple trend-noise specifications.

Table A.2: Diagnostics for Series A: 'Petrol_Cons_Ind_SA'

| Trend | Q(12) | Q(24) | Q(36) | Eq SE | AIC | SIC | R^2_d |
|---|---|---|---|---|---|---|---|
| Standard (Order 1) | 37.77 | 85.01 | 114 | 0.04099 | -3071.91 | -2881.48 | 0.264 |
| Standard (Order 2) | 60.28 | 101.6 | 124.4 | 0.04279 | -2881.27 | -2690.84 | 0.198 |
| Standard (Order 3) | 129.8 | 184.6 | 208 | 0.04458 | -2720.14 | -2529.71 | 0.129 |
| Standard (Order 4) | 134.8 | 200.9 | 219.2 | 0.04549 | -2640.58 | -2450.15 | 0.0933 |
| Standard (Order 5) | 141.6 | 219.3 | 239.6 | 0.04632 | -2563.23 | -2372.80 | 0.0603 |

Note: See notes to table A.1.

Table A.3: Maximum likelihood estimates of parameters for Series: 'Petrol_Cons_Ind_SA'

| Trend | ZetaVar | EpsVar | Q_Zeta | BetaMean | Phi |
|---|---|---|---|---|---|
| Canonical (Order 1) | 8.305e-005 | 0.00105 | 0.07912 | 0.000109 | N/A |
| Canonical (Order 2) | 1.574e-006 | 0.001366 | 0.001153 | 0.0001206 | 0.95 |
| Canonical (Order 3) | 1.332e-008 | 0.001529 | 8.713e-006 | 0.0003239 | 0.95 |
| Canonical (Order 4) | 5.731e-011 | 0.001666 | 3.44e-008 | -0.0009026 | 0.95 |
| Canonical (Order 5) | 2.458e-013 | 0.001746 | 1.407e-010 | -0.003382 | 0.95 |

Note: See note to table 1. Zeta now refers to the trend disturbance for the canonical trend model.

Table A.4: Estimated Trend and RMSE values for Series: 'Petrol_Cons_Ind_SA'

| Trend | Trend-Mid | RMSE-Mid | RMSE-Pct | Trend-End | RMSE-End | RMSE-Pct |
|---|---|---|---|---|---|---|
| Canonical (Order 1) | 8.447 | 0.0152 | 0.1797% | 8.4690 | 0.0203 | 0.239% |
| Canonical (Order 2) | 8.443 | 0.0112 | 0.1328% | 8.4665 | 0.0195 | 0.230% |
| Canonical (Order 3) | 8.435 | 0.0096 | 0.1142% | 8.4698 | 0.0193 | 0.228% |
| Canonical (Order 4) | 8.433 | 0.0087 | 0.1028% | 8.4705 | 0.0189 | 0.224% |
| Canonical (Order 5) | 8.435 | 0.0081 | 0.0961% | 8.4696 | 0.0188 | 0.222% |

Note: See notes to table 2.

Table A.5: Diagnostics for Series: 'Petrol_Cons_Ind_SA'

| Trend | Q(12) | Q(24) | Q(36) | Eq SE | AIC | SIC | R^2_d |
|---|---|---|---|---|---|---|---|
| Canonical (Order 1) | 45.47 | 96.31 | 127.9 | 0.0415 | -1896.43 | -1896.43 | 0.2451 |
| Canonical (Order 2) | 108.8 | 164.6 | 200.6 | 0.0435 | -1845.42 | -1845.42 | 0.172 |
| Canonical (Order 3) | 132.1 | 196.6 | 227 | 0.045 | -1807.75 | -1807.75 | 0.1135 |
| Canonical (Order 4) | 148.9 | 249.2 | 273.7 | 0.0461 | -1781.69 | -1781.69 | 0.06969 |
| Canonical (Order 5) | 163.9 | 288.5 | 309.6 | 0.0468 | -1765.41 | -1765.41 | 0.04044 |

# Appendix B   Additional Figures

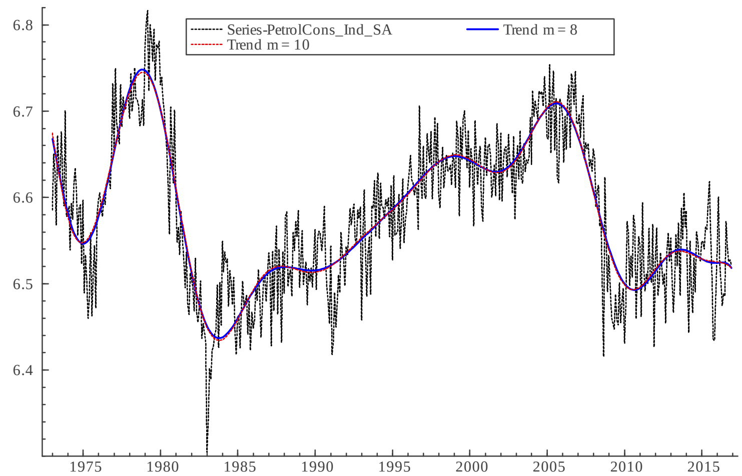


Figure B.1: Estimated trend in Petroleum Consumption for univariate eighth and tenth order standard model.

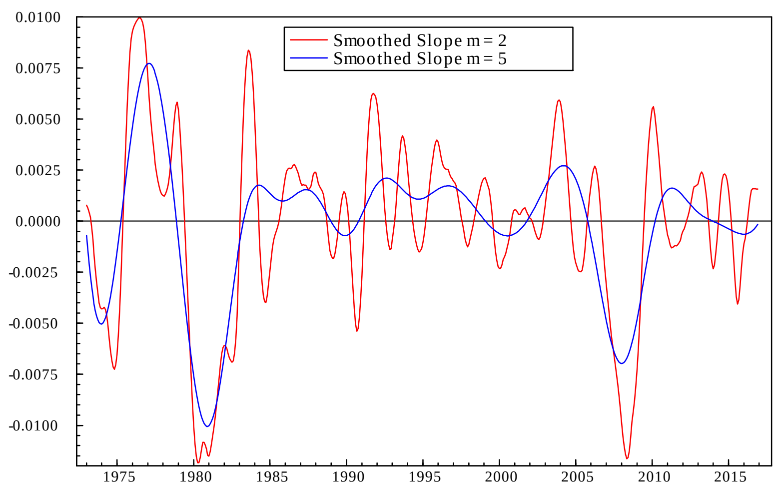


Figure B.2: Estimated slope of signal in Petroleum Consumption for univariate second- and fifth-order standard models.

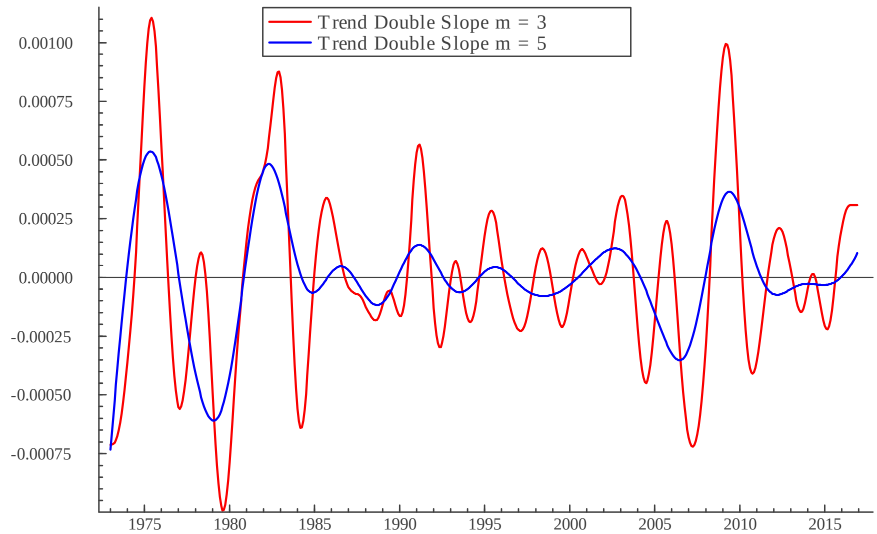


Figure B.3: Estimate rate of change of slope of signal in Petroleum Consumption for univariate third- and fifth-order standard models.

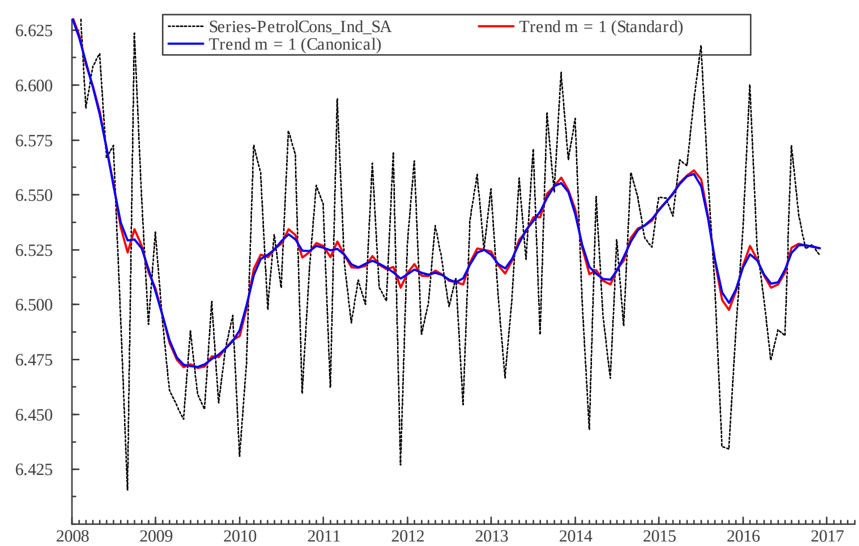


Figure B.4: Estimated trend in Petroleum Consumption for standard and canonical specificiations, based on univariate first-order models.

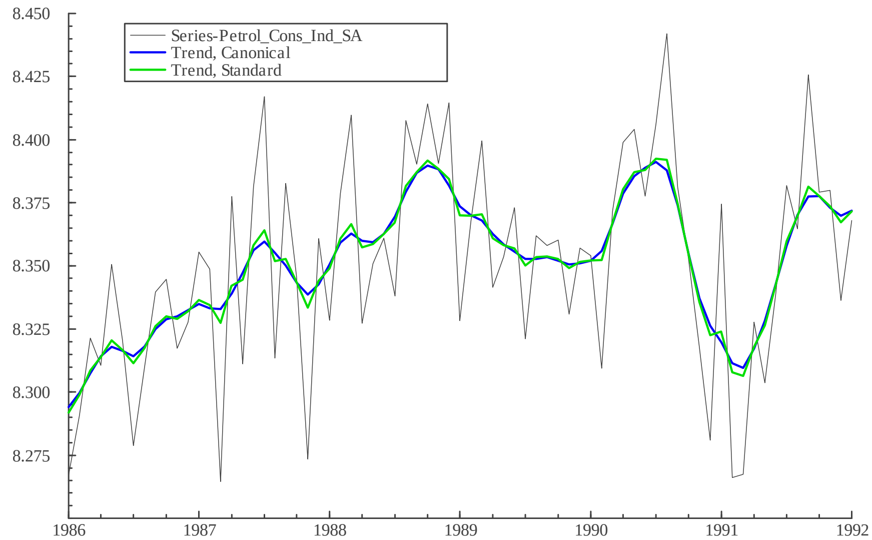


Figure B.5: Estimated trend in Petroleum Consumption for univariate first-order standard and canonical models for sample period from January 1986 to December 1991.

# Appendix C Model-based seasonal adjustment of petroleum-related time series

The original data on U.S. Petroleum Consumption (seasonally unadjusted and in logarithms) are shown in figure C.1. For the application, this data series was seasonally adjusted via a model-based approach. In particular, univariate models of the form

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \tag{C.1}$$

were used, where the model for the stochastic seasonal component $\gamma_t$ is specified as:

$$\begin{aligned} \gamma_t &= \textstyle\sum_{j=1}^{[s/2]} \gamma_{j,t}, \\ \gamma_{j,t+1} &= \gamma_{j,t} \cos \lambda_j + \gamma_{j,t}^* \sin \lambda_j + \omega_{j,t}, \\ \gamma_{j,t+1}^* &= \gamma_{j,t} \cos \lambda_j - \gamma_{j,t}^* \sin \lambda_j + \omega_{j,t}^*, \\ \omega_{j,t}, \omega_{j,t}^* &\sim i.i.d.\ N(0, \sigma_\omega^2). \end{aligned} \tag{C.2}$$

This specification starts with a Fourier decomposition and then makes each component cycle stochastic. The seasonal frequencies (nonstationary) stochastic cycles are given by

$$\lambda_j = 2\pi j/s, \quad j = 1, ..., s/2-1,$$
$$\lambda_j = \pi, \quad j = s/2.$$

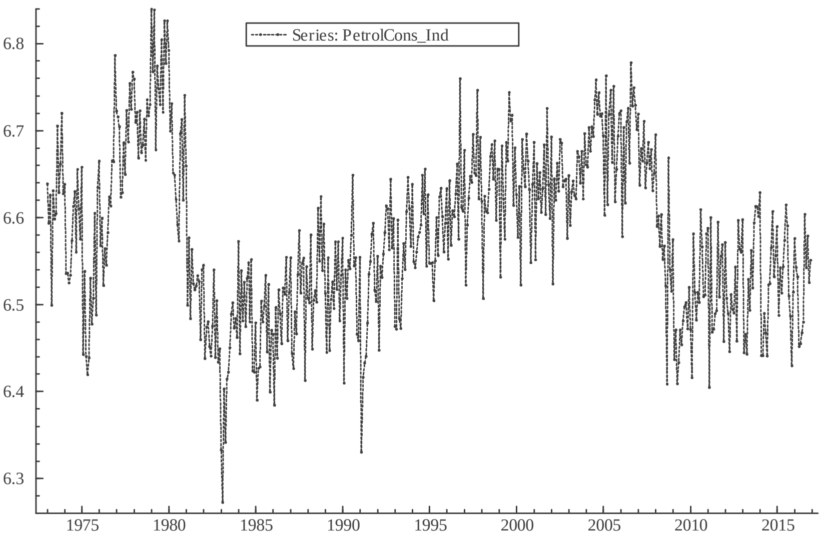


Figure C.1: Time Series of U.S. Petroleum Consumption (Industrial Sector), Non-Seasonally-Adjusted and in logarithms.

Given the constant variance $\sigma^2_\omega$ for all $j$, the model (C.2) is referred to as TRIG-1. This contrasts with a more general model proposed by Harvey (1989), in which $\text{Var}(\omega_{j,t})$ is distinct for each $j = 1, ..., s/2$ (called "TRIG-6"). Here we focus on the TRIG-1 model, which is more parsimonious. The white noise disturbances $\eta_t$, $\zeta_t$, $\omega_{j,t}$, and $\omega^*_{j,t}$ are assumed uncorrelated with one another across all time points, and also with the irregular.

The extracted seasonal component, denoted by $\widehat{\gamma}_t$, is indicated in Figure C.2. The seasonally adjusted estimates are then given by $y_t - \widehat{\gamma}_t$,

# Appendix D Proof of Propositions

**Proposition 1** *For $\phi \approx 1$, the cutoff frequency, at which the gain equals one-half, is approximately*

$$\lambda_{1/2} = 2 \arcsin(2^{-1/2}[q_\zeta / r(m, \phi)]^{1/2m}), \tag{D.1}$$

*where $r(m, \phi) = 1 + (m-1)(-1+\phi)$.*

**Proof.**

We seek approximate solutions for $\phi \approx 1$. Consider the function

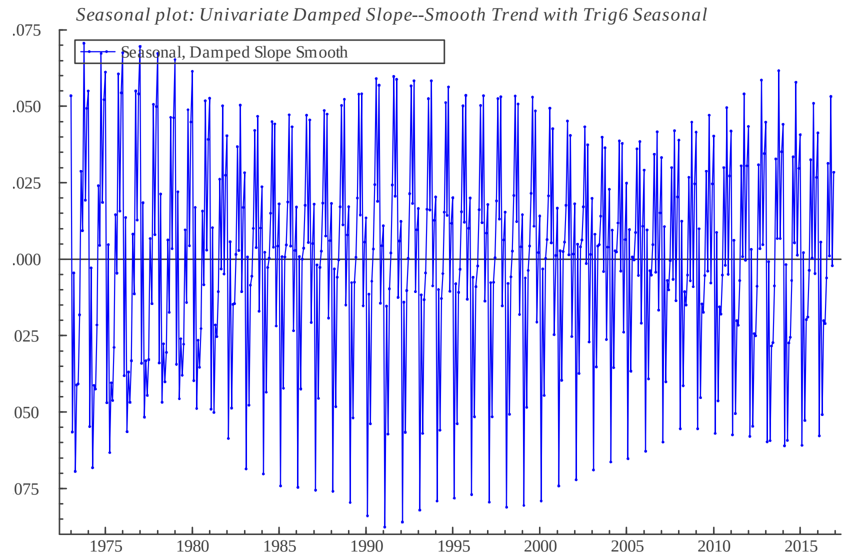


Figure C.2: Extracted Seasonal component in U.S. Petroleum Consumption (Industrial Sector).

$$f(\phi) = (2 - 2\cos\lambda)(1 + \phi^2 - 2\phi\cos\lambda)^{m-1}.$$

In the vicinity of $\phi = 1$ (with $\phi$ strictly less than one), with sufficient regularity conditions on $f(\phi)$,

$$f(\phi) = f(1) + f^{'}(1)[\phi - 1].$$

Furthermore,

$$f^{'}(\phi) = (m-1)(2 - 2\cos\lambda)(1 + \phi^2 - 2\phi\cos\lambda)^{m-2}\left(2\phi - 2\cos\lambda\right),$$

so that

$$f^{'}(1) = (m-1)(2 - 2\cos\lambda)^m.$$

Now,

$$\begin{aligned} f(\phi) &\approx (2 - 2\cos\lambda)^m + (m-1)(2 - 2\cos\lambda)^m[\phi - 1] \\ f(\phi) &\approx [1 + (m-1)(\phi - 1)](2 - 2\cos\lambda)^m \end{aligned}$$

Let $r(m,\phi) = [1 + (m-1)(\phi - 1)]$. Then

$$q_\zeta = r(m,\phi)(2 - 2\cos\lambda)^m,$$

or, using the cosine-addition formula for $\cos(\lambda/2 + \lambda/2)$,

$$q_\zeta = r(m,\phi)[4\sin^2(\lambda/2)]^m.$$

Solving this equation for $\lambda$ yields (D.1).

■

**Proposition 2** *The elements of the gain matrix for the bivariate Butterworth filter are given by*

$$\begin{aligned}
G_A(\lambda) &= h(\lambda)[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B} - \nu_{\zeta,AB}\nu_{\varepsilon,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}] \\
G_{AB}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,A}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B} + \nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,A}] \\
G_{BA}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,B}\nu_{\varepsilon,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B} + \nu_{\zeta,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma^2_{\varepsilon,B}] \\
G_B(\lambda) &= h(\lambda)[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,B}\sigma^2_{\varepsilon,A} - \nu_{\zeta,AB}\nu_{\varepsilon,AB}\sigma_{\zeta,A}\sigma_{\zeta,B}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}]
\end{aligned}$$

*where*

$$\begin{aligned}
h(\lambda) &= [f_\mu(\lambda)/R(\lambda)] = 1/[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + (1-\nu^2_{\varepsilon,AB})\sigma^2_{\varepsilon,A}\sigma^2_{\varepsilon,B}/f_\mu(\lambda) + (\sigma^2_{\varepsilon,A}\sigma^2_{\zeta,B} + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B}) \\
&\quad -2\nu_{\varepsilon,AB}\nu_{\zeta,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}\sigma_{\zeta,A}\sigma_{\zeta,B}]
\end{aligned}$$

**Proof.**

$$[\Sigma_\zeta f_\mu(\lambda) + \Sigma_\varepsilon] = \begin{bmatrix} \sigma^2_{\zeta,A}f_\mu(\lambda) + \sigma^2_{\varepsilon,A} & \sigma_{\zeta,AB}f_\mu(\lambda) + \sigma_{\varepsilon,AB} \\ \sigma_{\zeta,AB}f_\mu(\lambda) + \sigma_{\varepsilon,AB} & \sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\varepsilon,B} \end{bmatrix}$$

Then we seek the inverse of the above matrix.

$$[\Sigma_\zeta f_\mu(\lambda) + \Sigma_\varepsilon]^{-1} = [1/R(\lambda)] \begin{bmatrix} \sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\varepsilon,B} & -\sigma_{\zeta,AB}f_\mu(\lambda) - \sigma_{\varepsilon,AB} \\ -\sigma_{\zeta,AB}f_\mu(\lambda) - \sigma_{\varepsilon,AB} & \sigma^2_{\zeta,A}f_\mu(\lambda) + \sigma^2_{\varepsilon,A} \end{bmatrix}$$

Computing the determinant, or the multiplier's reciprocal, gives the expression

$$\begin{aligned}
R(\lambda) &= (1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}[f_\mu(\lambda)]^2 + (1-\nu^2_{\varepsilon,AB})\sigma^2_{\varepsilon,A}\sigma^2_{\varepsilon,B} + (\sigma^2_{\varepsilon,A}\sigma^2_{\zeta,B} + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B})f_\mu(\lambda) \\
&\quad -2\nu_{\varepsilon,AB}\nu_{\zeta,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}\sigma_{\zeta,A}\sigma_{\zeta,B}f_\mu(\lambda),
\end{aligned}$$

so the gain matrix is

$$\Sigma_\zeta f_\mu(\lambda)[\Sigma_\zeta f_\mu(\lambda) + \Sigma_\varepsilon]^{-1} = [f_\mu(\lambda)/R(\lambda)]\Sigma_\zeta \begin{bmatrix} \sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\varepsilon,B} & -\sigma_{\zeta,AB}f_\mu(\lambda) - \sigma_{\varepsilon,AB} \\ -\sigma_{\zeta,AB}f_\mu(\lambda) - \sigma_{\varepsilon,AB} & \sigma^2_{\zeta,A}f_\mu(\lambda) + \sigma^2_{\varepsilon,A} \end{bmatrix}.$$

Let

$$
\begin{aligned}
h(\lambda) &= [f_\mu(\lambda)/R(\lambda)] = 1/[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + (1-\nu^2_{\varepsilon,AB})\sigma^2_{\varepsilon,A}\sigma^2_{\varepsilon,B}/f_\mu(\lambda) + (\sigma^2_{\varepsilon,A}\sigma^2_{\zeta,B} + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B}) \\
&\quad -2\nu_{\varepsilon,AB}\nu_{\zeta,AB}\sigma_{\varepsilon,A}\sigma_{\varepsilon,B}\sigma_{\zeta,A}\sigma_{\zeta,B}]
\end{aligned}
$$

so that

$$
\Sigma_\zeta f_\mu(\lambda)[\Sigma_\zeta f_\mu(\lambda)+\Sigma_\varepsilon]^{-1} = h(\lambda)\begin{bmatrix} \sigma^2_{\zeta,A} & \sigma_{\zeta,AB} \\ \sigma_{\zeta,AB} & \sigma^2_{\zeta,B} \end{bmatrix}\begin{bmatrix} \sigma^2_{\zeta,B}f_\mu(\lambda)+\sigma^2_{\varepsilon,B} & -\sigma_{\zeta,AB}f_\mu(\lambda)-\sigma_{\varepsilon,AB} \\ -\sigma_{\zeta,AB}f_\mu(\lambda)-\sigma_{\varepsilon,AB} & \sigma^2_{\zeta,A}f_\mu(\lambda)+\sigma^2_{\varepsilon,A} \end{bmatrix}.
$$

So, the elements of the gain matrix are

$$
\begin{aligned}
\textit{element } 1{,}1 &: \quad \sigma^2_{\zeta,A}(\sigma^2_{\zeta,B}f_\mu(\lambda)+\sigma^2_{\varepsilon,B}) + \sigma_{\zeta,AB}(-\sigma_{\zeta,AB}f_\mu(\lambda)-\sigma_{\varepsilon,AB}) = \\
&\quad \sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B} - (\sigma_{\zeta,AB})^2 f_\mu(\lambda) - \sigma_{\zeta,AB}\sigma_{\varepsilon,AB} \\
G_A(\lambda) &= h(\lambda)[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,A}\sigma^2_{\varepsilon,B} - \sigma_{\zeta,AB}\sigma_{\varepsilon,AB}] \\
\textit{element } 1{,}2&: \quad \sigma^2_{\zeta,A}(-\sigma_{\zeta,AB}f_\mu(\lambda)-\sigma_{\varepsilon,AB}) + \sigma_{\zeta,AB}(\sigma^2_{\zeta,A}f_\mu(\lambda)+\sigma^2_{\varepsilon,A}) = \\
&\quad -\sigma^2_{\zeta,A}\sigma_{\zeta,AB}f_\mu(\lambda) - \sigma^2_{\zeta,A}\sigma_{\varepsilon,AB} + \sigma_{\zeta,AB}\sigma^2_{\zeta,A}f_\mu(\lambda) + \sigma_{\zeta,AB}\sigma^2_{\varepsilon,A} \\
G_{AB}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,A}\sigma_{\varepsilon,AB} + \sigma_{\zeta,AB}\sigma^2_{\varepsilon,A}] \\
\textit{element } 2{,}1 &: \quad \sigma^2_{\zeta,B}(-\sigma_{\zeta,AB}f_\mu(\lambda)-\sigma_{\varepsilon,AB}) + \sigma_{\zeta,AB}(\sigma^2_{\zeta,B}f_\mu(\lambda)+\sigma^2_{\varepsilon,B}) = \\
&\quad -\sigma^2_{\zeta,B}\sigma_{\zeta,AB}f_\mu(\lambda) - \sigma^2_{\zeta,B}\sigma_{\varepsilon,AB} + \sigma_{\zeta,AB}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma_{\zeta,AB}\sigma^2_{\varepsilon,B} \\
G_{BA}(\lambda) &= h(\lambda)[-\sigma^2_{\zeta,B}\sigma_{\varepsilon,AB} + \sigma_{\zeta,AB}\sigma^2_{\varepsilon,B}] \\
\textit{element } 2{,}2 &: \quad \sigma_{\zeta,AB}(-\sigma_{\zeta,AB}f_\mu(\lambda)-\sigma_{\varepsilon,AB}) + \sigma^2_{\zeta,B}(\sigma^2_{\zeta,A}f_\mu(\lambda)+\sigma^2_{\varepsilon,A}) = \\
&\quad \sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,B}\sigma^2_{\varepsilon,A} - (\sigma_{\zeta,AB})^2 f_\mu(\lambda) - \sigma_{\zeta,AB}\sigma_{\varepsilon,AB} \\
G_B(\lambda) &= h(\lambda)[(1-\nu^2_{\zeta,AB})\sigma^2_{\zeta,A}\sigma^2_{\zeta,B}f_\mu(\lambda) + \sigma^2_{\zeta,B}\sigma^2_{\varepsilon,A} - \sigma_{\zeta,AB}\sigma_{\varepsilon,AB}]
\end{aligned}
$$

■

# Appendix E Filter Formula in general case

When there are common trends, or co-integration, $\Sigma_\zeta$ is of reduced rank; to handle this possibility a strategy is needed to guarantee the existence of the multivariate WK filter. First, we re-express the filter formula to be in terms of $\boldsymbol{Q}$ rather than $\boldsymbol{Q}^{-1}$.

$$LP(L) = \boldsymbol{Q}\boldsymbol{Q}^{-1}\big(I_N + [(1-L)(1-L^{-1})]^m \boldsymbol{Q}^{-1}\big)^{-1}. \tag{E.1}$$

$$LP(L) = \boldsymbol{Q}\big[\big(I_N + [(1-L)(1-L^{-1})]^m \boldsymbol{Q}^{-1}\big)\,\boldsymbol{Q}\big]^{-1}. \tag{E.2}$$

$$LP(L) = \boldsymbol{Q}[\boldsymbol{Q} + [(1-L)(1-L^{-1})]^m \boldsymbol{Q}^{-1}\boldsymbol{Q})]^{-1}. \tag{E.3}$$

$$LP(L) = \boldsymbol{Q}[\boldsymbol{Q} + [(1-L)(1-L^{-1})]^m I_N]^{-1}. \tag{E.4}$$

The frequency response is

$$LP(\lambda) = \boldsymbol{Q}(\boldsymbol{Q} + r(\lambda) I_N)^{-1}. \tag{E.5}$$

Now let

$$\boldsymbol{Q} = V\,\Lambda\,V^{-1}$$

be the eigen-decomposition of $\boldsymbol{Q}$. Therefore, now the decomposition is made of $\boldsymbol{Q}$ rather than $\boldsymbol{Q}^{-1}$. Furthermore, since $Q$ is the product of two symmetric matrices, one of which is positive definite ($\Sigma_{\boldsymbol{\varepsilon}}^{-1}$), its eigenvalues are real and non-negative. (Indeed, if $A$ and $B$ are symmetric matrices and $A$ is invertible, then the symmetric square root $A^{1/2}$ exists and is invertible, and hence $\det\{\lambda\, I_N - A\,B\} = \det\{\lambda\, I_N - A^{1/2}\,B\,A^{1/2}\}$, showing that the eigenvalues of $A\,B$ and $A^{1/2}\,B\,A^{1/2}$ are the same; the latter is symmetric, and hence has real non-negative eigenvalues.)

The filter can be written as

$$\begin{aligned}
LP(L) &= \boldsymbol{Q}(\boldsymbol{Q} + [(1-L)(1-L^{-1})]^m I_N)^{-1}. \\
&= V\,\Lambda\,V^{-1}(V\,\Lambda\,V^{-1} + h(L) I_N)^{-1}. \\
&= V\,\Lambda\,[(V\,\Lambda\,V^{-1} + h(L) I_N)V]^{-1}. \\
&= V\,\Lambda\,[(V\,\Lambda\,V^{-1}V + h(L)V]^{-1}. \\
&= V\,\Lambda\,[(V\,\Lambda + h(L)V]^{-1}. \\
&= V\,\Lambda\,[V\,(\Lambda + h(L) I_N)]^{-1}. \\
&= V\,\Lambda\,[(\Lambda + h(L) I_N]^{-1} V^{-1}.
\end{aligned}$$

where $h(L) = [(1-L)(1-L^{-1})]^m$.

The corresponding form of the frequency response is

$$LP(\lambda) = V\,\Lambda(\Lambda + r(\lambda) I_N)^{-1}\,V^{-1}. \tag{E.6}$$

For $\{X_t\}$ and $\{Y_t\}$ stationary $N$-variate processes related via $Y_t = \mathrm{LP}(L)\,X_t$, the spectral representation of the filter output is now

$$Y_t = \int_{-\pi}^{\pi} e^{i\lambda t}\,V\,\Lambda\,(\Lambda + r(\lambda)\,I_N)^{-1}\,V^{-1}\,dZ(\lambda),$$

$$Y_t = \int_{-\pi}^{\pi} e^{i\lambda t}\, V\, \Lambda\, (\Lambda + r(\lambda)\, I_N)^{-1}\; dZ^\star(\lambda),$$

The demand that the $j$th output has half the frequency content of the $j$th input has the form:

$$e_j'\, V\, \Lambda\, (\Lambda + r(\lambda)\, I_N)^{-1} = \frac{1}{2}\, e_j'\, V.$$

Taking the transpose,

$$(\Lambda + r(\lambda)\, I_N)^{-1'} \Lambda' V' e_j = \frac{1}{2}\, V' e_j,$$

which because $\Lambda$ is diagonal,

$$[(\Lambda + r(\lambda)\, I_N)^{-1} \Lambda\,] V' e_j = \frac{1}{2}\, V' e_j,$$

meaning that $V' e_j$ would be an eigen-vector of $(\Lambda + r(\lambda)\, I_N)^{-1}\Lambda$ with eigen-value 1/2; however, since the matrix is diagonal, the only eigen-vectors are actually unit vectors. Hence, the system has no solution unless $V$ is a permutation matrix.

However, as for the related trends case, we can allow for solving a less demanding problem: for each $1 \leq j \leq N$, we seek a $\lambda_j$ such that output has half the frequency content of the input when the increment process equals $e_j$ (corresponding to an impulse), i.e.,

$$\frac{1}{2}\, e_j'\, V\, e_j = e_j'\, V\, \Lambda\, (\Lambda + r(\lambda_j)\, I_N)^{-1}\, e_j.$$

This is solved as $r(\lambda_j) = \Lambda_{jj}$, i.e., set the eigen-values of $Q$ equal to $(2 - 2\cos(\lambda_j))^m$ and solve for $\lambda_j$.

In the case that $Q$ is invertible (if and only if $\Sigma_\zeta$ is invertible), then we can rewrite the filter in the earlier form where $S(\lambda)$ is diagonal now with $jj$th entry given by $r(\lambda)/\Lambda_{jj}$.