Schlicht, Ekkehart

Working Paper

Trend Extraction From Time Series With Structural Breaks

Munich Discussion Paper, No. 2007-17

Provided in Cooperation with:
University of Munich, Department of Economics

Suggested Citation: Schlicht, Ekkehart (2007) : Trend Extraction From Time Series With Structural Breaks, Munich Discussion Paper, No. 2007-17, Ludwig-Maximilians-Universität München, Volkswirtschaftliche Fakultät, München, https://doi.org/10.5282/ubm/epub.1926

This Version is available at:
https://hdl.handle.net/10419/104224

Ekkehart Schlicht:

Trend Extraction From Time Series With Structural Breaks

Munich Discussion Paper No. 2007-17

Department of Economics
University of Munich

Volkswirtschaftliche Fakultät
Ludwig-Maximilians-Universität München

Trend Extraction From Time Series With Structural Breaks

Ekkehart Schlicht

Abstract: Trend extraction from time series is often performed by using the filter proposed by Leser (1961), also known as the Hodrick-Prescott filter. A practical problem arises, however, when the time series contains structural breaks (such as produced by German unification for German time series, for instance). This note proposes a method for coping with this problem.

Keywords: Trend extraction, structural break, Hodrick-Prescott filter, Leser filter, spline, time-series, smoothing, interpolation.

Journal of Economic Literature Classification: C22, C32, C63, C14

1 Introduction

Trend extraction from time series is often performed by using the filter proposed by Leser (1961), also known as the Hodrick-Prescott filter, or HP-Filter. A practical problem arises, however, when the time series contains structural breaks (such as produced by German unification for German time series, for instance). This note proposes a method for coping with this problem.

2 The Leser Filter[1]

The idea proposed by Leser (1961) for the case where all data are available is to look for a trend $y \in \mathbb{R}^T$ such that deviation

$$u = x - y \tag{1}$$

is "small" and the trend is "smooth." The size of the deviation is measured by the sum of squared residuals $u'u$, and the smoothness of the trend is measured by the sum of squares of changes in the direction of the trend $v'v$ where the trend disturbances $v \in \mathbb{R}^{T-2}$ are defined as

$$v_t = ((y_t - y_{t-1}) - (y_{t-1} - y_{t-2})) \qquad t = 3, 4, \dots, T$$

or

$$v = P\,y \tag{2}$$

with

$$P := \begin{pmatrix} 1 & -2 & 1 & & & 0 \\ & 1 & -2 & 1 & & \\ & & \cdot & \cdot & \cdot & \\ 0 & & & 1 & -2 & 1 \end{pmatrix}$$

of order $(T-2) \times T$.

The decomposition of the original series x into trend y and and residual u is obtained by minimizing the weighted sum of squares

$$V = u'u + \alpha \cdot v'v = (x-y)'\,(x-y) + \alpha \cdot y'P'P\,y \tag{3}$$

[1] The formailization below follows Schlicht (1981).

with respect to y. This gives the first-order condition

$$(I_T + \alpha \cdot P'P)\, y = x. \tag{4}$$

As $(I + \alpha\, P'P)$ is positive definite, the second order condition is satisfied in any case.

Equation (4) has the unique solution

$$y = (I_T + \alpha\, P'P)^{-1}\, x \tag{5}$$

which defines the Leser-Filter. It associates a trend y with the time series x, depending on the smoothing parameter α.

From (3) and (5) we obtain

$$V = x' \left(I_T - (I_T + \alpha \cdot P'P)^{-1} \right) x \tag{6}$$

as the value of the criterion function (3).

3 *Structural Breaks*

A practical problem arises when the time series contains structural breaks. Equation (5) interprets these wrongly as disturbances. The obvious way to generalize the filter in order to cope with this problem is to look for a smooth trend by introducing dummy variables to capture these effects, and to select these variables such that the criterion function (3) is minimized. This can be done as follows.

Consider a time series $x \in \mathbb{R}^T$ with $m \leq T - 2$ structural breaks. The break points are indicated by a vector $m \in \mathbb{R}^T$ with components $m_t = 1$ for all $t \in \{1, 2, ...T\}$ where a break point occurs, and $m_t = 0$ otherwise. Denote the break points by $t_1, t_2, .., t_m$. Define the $(T \times m)$-matrix with elements $b_{i,j} = 0$ for $i < t_j$ and $b_{i,j} = 1$ for $i \geq t_j$. For $T = 5$ and break points at $t_1 = 2$ and $t_2 = 4$ we would have, for example

$$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Define further the vector of dummies dor the break points $d \in \mathbb{R}^m$, where d_i gives the dummy for the i-th break point.

The adjusted vector of observations is

$$x^* = x + Bd. \tag{7}$$

It is a function of the values assumed for the dummies d. These values can now easily be determined by replacing x by x^* in (6), and minimizing this expression. Thus we obtain the positive definite quadratic form

$$V = (z' + s'B') \left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right) (z + Bd) \tag{8}$$

that is to be minimized with respect to s.The first-order conditions for a minimum is

$$\frac{\partial V}{\partial s} = 2B' \left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right) (z + Bs) = 0. \tag{9}$$

and the second-order condition is that $B' \left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right) B$ is positive definite. As

$$(I_T + \alpha P'P)^{-1} = I_T - \alpha P'P + (\alpha P'P)^2 - (\alpha P'P)^3 + ...$$

we can write

$$\left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right) = \alpha P' \left(I_{T-2} + \alpha \cdot PP'\right)^{-1} P.$$

Hence $\left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right)$ is non-negative definite of rank $T - 2$ and $B' \left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right) B$ has full rank and is positive definite. Therefore equation (9) defines the unique maximizing choice of the dummy terms as

$$d^* = -\left(B' \left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right) B\right)^{-1} B' \left(I_T - (I_T + \alpha \cdot P'P)^{-1}\right) z. \tag{10}$$

The replenished time series is obtained now by inserting (10) into (7).

$$x^* = z + Bd^*$$

and the trend is obtained by using the replenished series x^* instead of the original series x

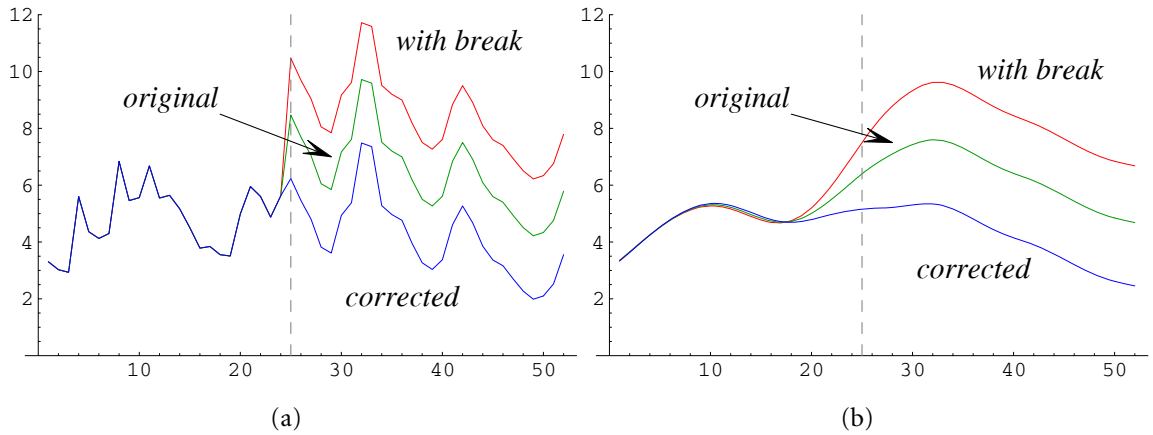

Figure 1: (a) The original US unemployment rate 1959-2002 has been augmented by adding, beginning with period 25, two percentage points to the original series. The corrected series overcorrects the structural break in this case. (b) The smoothed series reproduce this pattern. (Smothing constant is α =100, data from the Bureau of Labor Statistics.)

in (5):

$$y^* = (I_T + \alpha\, P'P)^{-1}\,(z + Bs^*)\,.$$

This gives the trend the time series, adjusted for structural breaks.

4 *An Example*[1]

As an example, consider the time series of US unemployment (Figure 1). Beginning with period 25, a structural break has been introduced by adding 2 percentage points to the original time series. The correction obtained by the method sketched above overcorrects this break by subtracting 4.2 percentage points. The corrected trend estimation is overcorrected as well. As can be seen, the correction produces a smoother trend than the original one, as is implied by the logic of the method. A manual correction would look not very much different, or would look even worse if the adjustment is made such that the adjacent data points 24 and 25 are made to have identical values. (The correction would have been -4.8 rather than -4.2 in this case.)

[1] All computations done with the package by Ludsteck (2004).

5 *Concluding Comments*

The example illustrates the functioning, as well as the problematic, of introducing dummies, as these will not only correct for structural breaks, but will also mask changes in the underlying trend. Yet there seems no better way available to do such corrections. If the Leser method is interpreted stochastically as in Schlicht (2005), a theoretical justification is that it can be interpreted as giving Maximum-Likelihood estimates for the dummies that capure structural breaks.

References

Leser, C. E. V. 1961, "A Simple Method of Trend Construction," *Journal of the Royal Statistical Society. Series B (Methodological)*, 23, 91–107.

Ludsteck, J. 2004, "HPFilter," Mathematica package, Wolfram Research, online at *http://library.wolfram.com/infocenter/MathSource/Mathematics/ProbabilityStatistics/*.

Schlicht, E. 1981, "A Seasonal Adjustment Principle and a Seasonal Adjustment Method Derived from This Principle," *Journal of the American Statistical Association*, 76(374), 374–78, online at *http://www.semverteilung.vwl.uni-muenchen.de/mitarbeiter/es/paper/schlicht-seasonal_adjustment.pdf*.

— 2005, "Estimating the Smoothing Constant in the So-Called Hodrick-Prescott Filter," *Journal of the Japan Statistical Society*, 35, 1–19, online at *http://www.scipress.org/journals/jjss/pdf/3501/35010099.pdf*.