1 HP97

The Hodrick and Prescott (1997, HP) Filter can be expressed in State Space Form (SSF) using the following Unobserved Component (UC) model structure:

$$
\begin{aligned}
y_t &= y_t^* + y_t^c && \text{(1a)}\\
\Delta^2 y_t^* &= \varepsilon_{1t} && \text{(1b)}\\
y_t^c &= \psi \varepsilon_{2t}, && \text{(1c)}
\end{aligned}
$$

where y_t is (100 times) the log of GDP, and the shocks $\{\varepsilon_{it}\}_{i=1}^2$ are mutually uncorrelated $i.i.d.$ $N(0,1)$. The standard deviation ψ is the (square root of the) smoothing parameter, commonly set to 40 for quarterly macroeconomic data, implying a value of 'λ' of 1600.

The '*numbered shock*' to '*named shock*' mapping is:

$$
\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_t^* \\ \varepsilon_t^c \end{bmatrix}, \tag{2}
$$

where ε_t^* is the trend (or permanent) shock and ε_t^c is the cycle (or transitory) shock.

2 Shock recovery

2.1 State Space Models with lagged states

Kurz (2018) adopts the following general SSF with lagged states in the measurement:

$$
\begin{aligned}
\text{Measurement}: \quad Z_t &= D_1 X_t + D_2 X_{t-1} + R\varepsilon_t && \text{(3a)}\\
\text{State}: \quad X_t &= \Phi X_{t-1} + Q\varepsilon_t, && \text{(3b)}
\end{aligned}
$$

where $\varepsilon_t \sim MN(0_m, I_m)$, D_1, D_2, Φ, R are Q are conformable system matrices, Z_t the observed variable and X_t the latent state variable, and m is the number of shocks $\{\varepsilon_{it}\}_{i=1}^2$.

2.2 HP97 in '*shock recovery*' SSF

To assess shock recovery, write the model in (1) in '*shock recovery*' SSF by collecting all observable variables in Z_t and all shocks (and other latent state variables) in X_t. Differencing y_t and y_t^c twice, and re-arranging the relations in (1) then yields:

$$
\begin{aligned}
\Delta^2 y_t &= \Delta^2 y_t^* + \Delta^2 y_t^c \\
&= \varepsilon_{1t} + \psi \Delta^2 \varepsilon_{2t} \\
&= \varepsilon_{1t} + \psi \varepsilon_{2t} - 2\psi \varepsilon_{2t-1} + \psi \varepsilon_{2t-2}, && \text{(4)}
\end{aligned}
$$

where $\Delta^2 y_t$ is the only observed variable.

The Measurement and State equations of the '*shock recovery*' SSF corresponding to the relations in (4) are then given by:

$$\begin{aligned}
\text{Measurement}: \quad & Z_t = D_1 X_t + D_2 X_{t-1} + R\varepsilon_t \\
\Leftrightarrow \Delta^2 y_t & = \varepsilon_{1t} + \psi\varepsilon_{2t} - 2\psi\varepsilon_{2t-1} + \psi\varepsilon_{2t-2}
\end{aligned} \tag{5a}$$

$$Z_t = \underbrace{\begin{bmatrix} 1 & \psi & 0 \end{bmatrix}}_{D_1} \underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{2t-1} \end{bmatrix}}_{X_t} + \underbrace{\begin{bmatrix} 0 & -2\psi & \psi \end{bmatrix}}_{D_2} \underbrace{\begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \\ \varepsilon_{2t-2} \end{bmatrix}}_{X_{t-1}} + \underbrace{\begin{bmatrix} 0 & 0 \end{bmatrix}}_{R} \underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}}_{\varepsilon_t} \tag{5b}$$

$$\text{State}: \quad X_t = \Phi X_{t-1} + Q\varepsilon_t,$$

$$\underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{2t-1} \end{bmatrix}}_{X_t} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \\ \varepsilon_{2t-2} \end{bmatrix}}_{X_{t-1}} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}}_{\varepsilon_t}. \tag{5c}$$

2.3 Shock recovery

The diagonal entries of the steady-state variance/covariance matrix of the smoothed and filtered states X_t denoted by $P^*_{t|T}$ and $P^*_{t|t}$, respectively, are:

Shocks	$\text{diag}(P^*_{t\|T})$	$\text{diag}(P^*_{t\|t})$
ε_{1t}	0.9439	0.9995
ε_{2t}	0.0561	0.2006

, (6)

indicating that the trend (or permanent) shock $\varepsilon_{1t} = \varepsilon^*_t$ cannot be recovered ($P^* \approx 1$), while the cycle shock $\varepsilon_{2t} = \varepsilon^c_t$ appears to be recoverable ($P^* \approx 0$). In Figure 1, simulated states and estimated Kalman smoothed states are plotted for the two shocks of interest.

The correlation between the true (simulated) and estimated Kalman smoothed shocks can be analyzed by simply computing $\text{Corr}(X_t, \hat{X}_{t|T})$, where $X_t = \begin{bmatrix} \varepsilon_{1t} & \varepsilon_{2t} & \varepsilon_{2t-1} \end{bmatrix}'$ and $\hat{X}_{t|T} = E_T X_t = E_T \begin{bmatrix} \varepsilon_{1t} & \varepsilon_{2t} & \varepsilon_{2t-1} \end{bmatrix}'$, yielding (for the first two elements of X_t):

Shocks	$\text{Corr}(X_t, \hat{X}_{t\|T})$
ε_{1t}	0.2368
ε_{2t}	0.9713

. (7)

As can be seen from (7), the estimated permanent shock ε_{1t} is only weakly correlated (0.2368) with the true value. The estimated cyclical shock ε_{2t}, on the other hand, is highly correlated (0.9713) with the true value. These facts can also be seen from Figure 1 below.

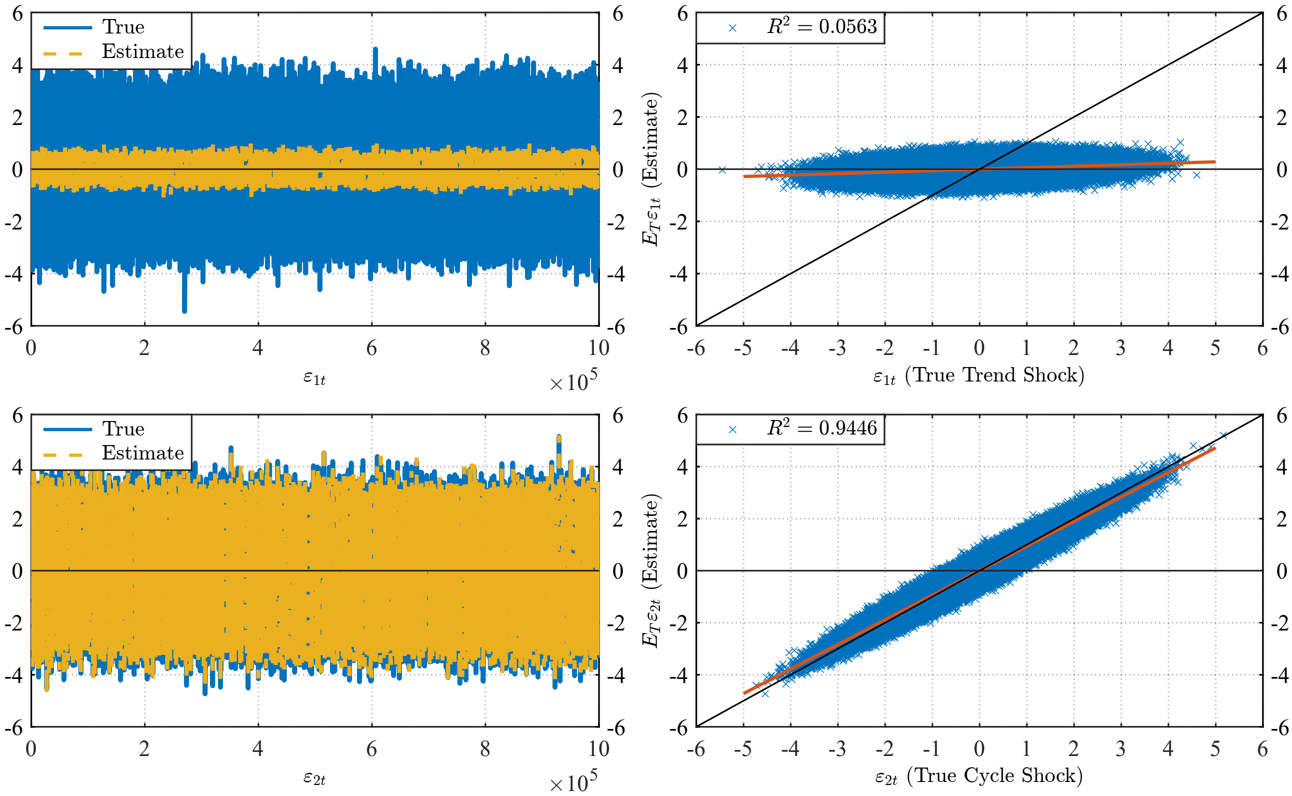

Figure 1: Comparison of true shocks and Kalman Smoothed estimates $\varepsilon_{t|T}$.

2.4 Shock Identities

Kalman Filter estimates of the permanent and transitory shocks ε_{1t} and ε_{2t} are linked by the identity:

$$E_t\varepsilon_{2t} = \psi E_t\varepsilon_{1t}, \tag{8}$$

and the corresponding Kalman Smoother estimates are linked by the *dynamic* identity:

$$\Delta^2 E_T\varepsilon_{1t} = \frac{1}{\psi} E_T\varepsilon_{2t-2}. \tag{9}$$

The filtered and smoothed estimates of the contemporaneous correlations are, respectively:

$\mathrm{Corr}(\hat{X}_{t|t}, \hat{X}_{t|t})$		
Shocks	ε_{1t}	ε_{2t}
ε_{1t}	1.0000	1.0000
ε_{2t}	1.0000	1.0000

and

$\mathrm{Corr}(\hat{X}_{t|T}, \hat{X}_{t|T})$		
Shocks	ε_{1t}	ε_{2t}
ε_{1t}	1.0000	-0.1907
ε_{2t}	-0.1907	1.0000

.

Note that ε_{1t} and ε_{2t} were generated as *i.i.d.* $N(0,1)$ and mutually uncorrelated.

Running regressions corresponding to (8) and (9) obtained from the simulated data (without an intercept) yields regression coefficient of 40 and $0.025 = 1/40$ when the smoothing parameter was set to $\lambda = 1600 = 40^2$ in the simulation. The regression fit is perfect, yielding an R^2 of 1 and a residual sum of squares of exactly 0 (see Panels (a) and (b) in Table 1, respectively). The file `HP97.m` replicates the output summarized here.

```
==============================================================================================================
Smoother Identity - Dependent variable: Δ²ETε1(t)
--------------------------------------------------------------------------------------------------------------
NOTE:    Identity should be: Δ²ETε1(t) = 1/40 ETε2(t-2) (and not ETε2(t) as stated)
--------------------------------------------------------------------------------------------------------------
==============================================================================================================
Variable         Estimate      stderr(HAC)      t-stat(HAC)         p-value
--------------------------------------------------------------------------------------------------------------
ETε2(t-2)        0.025000         0.000000 86387073152284000.000000         0.000000
--------------------------------------------------------------------------------------------------------------
 R-squared                :                                         1.000000 |  No. of Regressors     :       1.000000
 Rbar-squared             :                                         1.000000 |  Plus Const.(if exist) :       1.000000
 SE of regression         :                                         0.000000 |  Mean(y)               :       0.000000
 Sum Squared Errors       :                                         0.000000 |  Stdev(y)              :       0.024297
 Log-likelihood           :                                  34684702.107497 |  AIC                   :     -72.207418
 F-statistic              :     74627264078180667389585510985891840.000000 |  AICc                  :     -72.207418
 Pr(F-statistic)          :                                         0.000000 |  BIC                   :     -72.207406
 No. of observations      :                                    999998.000000 |  HQIC                  :     -72.207415
 Std.err.MLE (div by T)   :                                         0.000000 |  DW-stat.              :       2.603240
 Include Pre-whitening    :                                         0.000000 |  HAC Trunct.Lag.       :      31.000000
==============================================================================================================
```

(a) Kalman Filter

```
==============================================================================================================
Filter Identity - Dependent variable: Etε2(t)
==============================================================================================================
Variable        Estimate      stderr(HAC)      t-stat(HAC)         p-value
--------------------------------------------------------------------------------------------------------------
Etε1(t)       40.000000         0.000000 2148831869013260288.000000         0.000000
--------------------------------------------------------------------------------------------------------------
 R-squared                :                                         1.000000 |  No. of Regressors     :       1.000000
 Rbar-squared             :                                         1.000000 |  Plus Const.(if exist) :       1.000000
 SE of regression         :                                         0.000000 |  Mean(y)               :       0.001377
 Sum Squared Errors       :                                         0.000000 |  Stdev(y)              :       0.894404
 Log-likelihood           :                                  34359382.750073 |  AIC                   :     -71.556641
 F-statistic              :  4617478401287020629803264248150753280.000000 |  AICc                  :     -71.556641
 Pr(F-statistic)          :                                         0.000000 |  BIC                   :     -71.556629
 No. of observations      :                                   1000000.000000 |  HQIC                  :     -71.556637
 Std.err.MLE (div by T)   :                                         0.000000 |  DW-stat.              :       1.999361
 Include Pre-whitening    :                                         0.000000 |  HAC Trunct.Lag.       :      31.000000
==============================================================================================================
```

(b) Kalman Smoother

Table 1: Shock identity regressions using Kalman filter and smoother estimates.

Since $\varepsilon_{1t} = \Delta^2 y_t^*$ and $\varepsilon_{2t} = \frac{1}{\psi} y_t^c$, this implies that the output from the standard HP–Filter will give the corresponding idenity to (9) as:

$$
\begin{aligned}
\Delta^2 \underbrace{E_T \varepsilon_{1t}}_{\Delta^2 y_t^*} &= \frac{1}{\psi} \underbrace{E_T \varepsilon_{2t-2}}_{\frac{1}{\psi} y_{t-2}^c} \\
\Leftrightarrow \qquad \Delta^4 y_t^* &= \frac{1}{\psi^2} y_{t-2}^c, \\
\Leftrightarrow \quad \Delta^4 \text{HP–trend}_t &= \frac{1}{\psi^2} \text{HP–cycle}_{t-2},
\end{aligned}
$$

so that a regression of the fourth differenced HP–trend on a twice lagged HP–cycle will give a coefficient estimate of $\frac{1}{40^2} = \frac{1}{1600} = 0.000625$. The HP-Filter appears to be able to recover the transitory cyclical shock $\varepsilon_{2t} = \varepsilon_t^c$, but not the permanent shock. HP-Filter estimates of ε_{1t} and ε_{2t} induce a negative correlation between trend-cycle shocks, ie., $\text{Corr}(E_T \varepsilon_{1t}, E_T \varepsilon_{2t}) = -0.1907$, despite being generated with zero correlation.