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:

$$y_t = y_t^* + y_t^c \tag{1a}$$
$$\Delta^2 y_t^* = \varepsilon_{1t} \tag{1b}$$
$$y_t^c = \phi\varepsilon_{2t}, \tag{1c}$$

where y_t is (100 times) the log of GDP, and ε_{1t} and ε_{2t} are $N(0,1)$. The standard deviation ϕ 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:

$$\text{Measurement}: \quad Z_t = D_1 X_t + D_2 X_{t-1} + R\varepsilon_t \tag{3a}$$
$$\text{State}: \quad X_t = AX_{t-1} + C\varepsilon_t, \tag{3b}$$

where $\varepsilon_t \sim MN(0, I_m)$, D_1, D_2, A, R are C are conformable system matrices, Z_t the observed variable and X_t the latent state variable.

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:

$$\begin{aligned} \Delta^2 y_t &= \Delta^2 y_t^* + \Delta^2 y_t^c \\ &= \varepsilon_{1t} + \phi\Delta^2\varepsilon_{2t} \\ &= \varepsilon_{1t} + \phi\varepsilon_{2t} - 2\phi\varepsilon_{2t-1} + \phi\varepsilon_{2t-2}, \end{aligned} \tag{4}$$

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:

$$\text{Measurement}: \quad Z_t = D_1 X_t + D_2 X_{t-1} + R\varepsilon_t$$

$$\Leftrightarrow \Delta^2 y_t = \varepsilon_{1t} + \phi\varepsilon_{2t} - 2\phi\varepsilon_{2t-1} + \phi\varepsilon_{2t-2} \tag{5a}$$

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

$$\text{State}: \quad X_t = AX_{t-1} + C\varepsilon_t,$$

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

2.3 Shock recovery

The diagonal 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:

	$P^*_{t\|T}$	$P^*_{t\|t}$
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, while it is likely that the cycle shock $\varepsilon_{2t} = \varepsilon^c_t$ can be recovered. 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 recovered Kalman smoothed shocks can be analyzed by simply computing $\text{Corr}(X_t, X_{t|T})$, where $X_t = E_t \begin{bmatrix} \varepsilon_{1t} & \varepsilon_{2t} & \varepsilon_{2t-1} \end{bmatrix}$ and $X_{t|T} = E_T X_t = E_T \begin{bmatrix} \varepsilon_{1t} & \varepsilon_{2t} & \varepsilon_{2t-1} \end{bmatrix}'$, yielding:

	$\text{Corr}(X_t, X_{t\|T})$
ε_{1t}	0.2368
ε_{2t}	0.9713

. (7)

The correlation of the estimated permanent shock ε_{1t} with the true (simulated) counterpart is only 0.2368, while the estimated cyclical shock ε_{2t} is highly correlated at 0.9713 with the true value. These facts can also be seen from Figure 1 below.

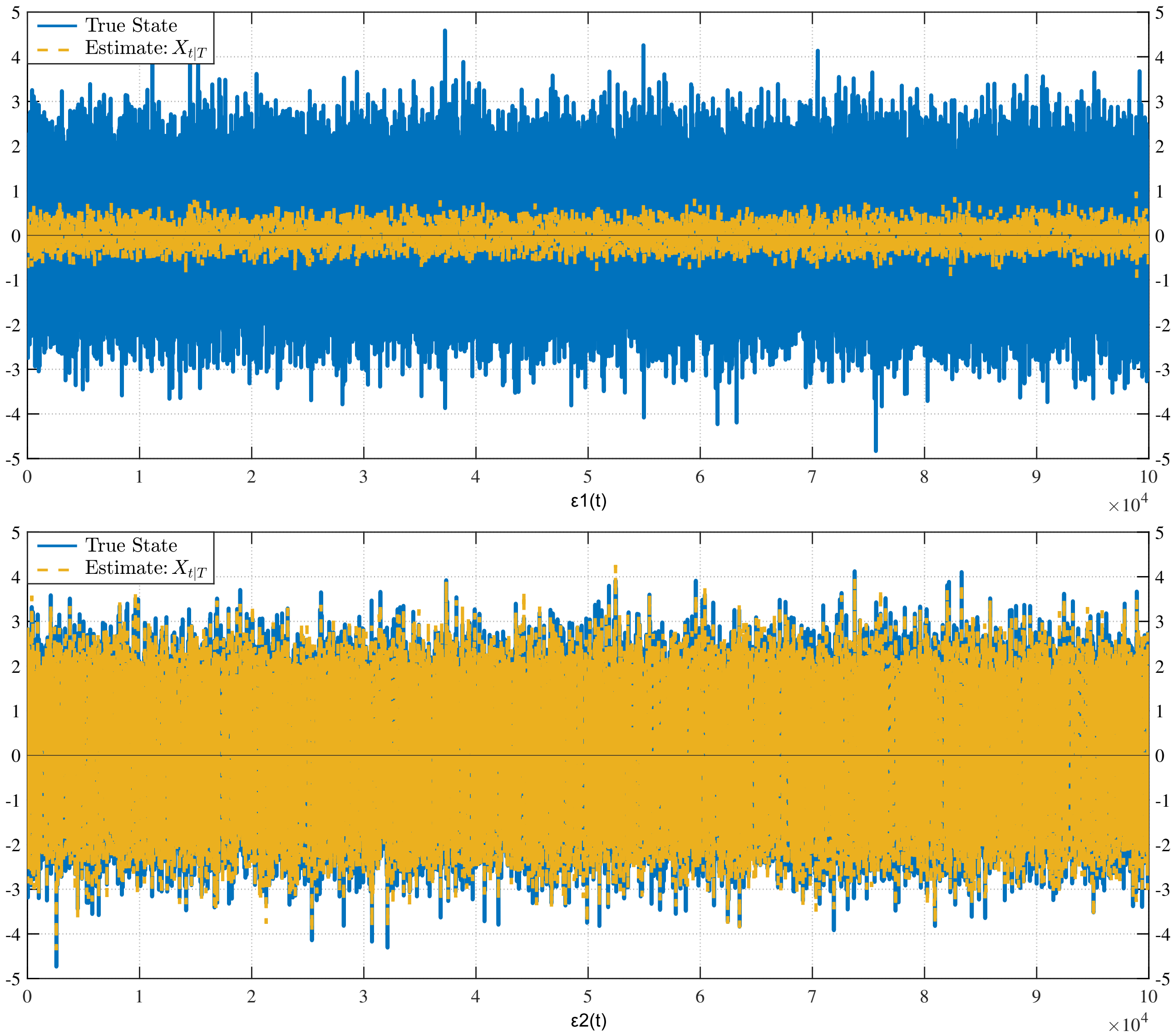

Figure 1: Comparison of true shocks and Kalman Smoothed estimates $\varepsilon_{t|T}$.

2.4 Shock Identities

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

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

and Kalman Smoother estimates are linked by the dynamic identity:

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

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

$\text{Corr}(X_{t|t}, X_{t|t})$

	ε_{1t}	ε_{2t}
ε_{1t}	1	1
ε_{2t}	1	1

and

$\text{Corr}(X_{t|T}, X_{t|T})$

	ε_{1t}	ε_{2t}
ε_{1t}	1	-0.1907
ε_{2t}	-0.1907	1

.

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

Since $\varepsilon_{1t} = \Delta^2 y_t^*$ and $\varepsilon_{2t} = \frac{1}{\phi} y_t^c$, this implies that the output from the standard HP–Filter will give the identity:

$$\Delta^4 y_t^* = \frac{1}{\phi} y_{t-2}^c, \quad \text{or alternatively:} \quad \Delta^4 \text{HP–trend}_t = \frac{1}{\phi} \text{HP–cycle}_{t-2}.$$

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.

```
================================================================================================================
Filter Identity. Dependent variable: Etε2(t)
================================================================================================================
Variable        Estimate      stderr(HAC)      t-stat(HAC)          p-value
----------------------------------------------------------------------------------------------------------------
Etε1(t)       40.000000         0.000000 619573092194908800.000000         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)                :         0.004701
 Sum Squared Errors        :                                           0.000000 |   Stdev(y)               :         0.898496
 Log-likelihood            :                                      355982.354905 |   AIC                    :       -74.034148
 F-statistic               :          3838708165719609501651610469788549120.000000 |   AICc                   :       -74.034148
 Pr(F-statistic)           :                                           0.000000 |   BIC                    :       -74.033427
 No. of observations       :                                       10000.000000 |   HQIC                   :       -74.033904
 Std.err.MLE (div by T) :                                              0.000000 |   DW-stat.               :         2.016162
 Include Pre-whitening  :                                              0.000000 |   HAC Trunct.Lag.        :        12.000000
================================================================================================================
```

(a) Kalman Filter

```
================================================================================================================
Smoother Identity. Dependent variable: Δ²ETε1(t)
================================================================================================================
Variable         Estimate      stderr(HAC)      t-stat(HAC)          p-value
----------------------------------------------------------------------------------------------------------------
ETε2(t-2)        0.025000         0.000000 20042737735416592.000000         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)                :     -0.000006
 Sum Squared Errors        :                                        0.000000 |   Stdev(y)               :      0.024396
 Log-likelihood            :                                   353018.214591 |   AIC                    :    -73.455443
 F-statistic               :       401711335930692148013508656627712.000000 |   AICc                   :    -73.455443
 Pr(F-statistic)           :                                        0.000000 |   BIC                    :    -73.454722
 No. of observations       :                                     9998.000000 |   HQIC                   :    -73.455199
 Std.err.MLE (div by T) :                                           0.000000 |   DW-stat.               :      3.659113
 Include Pre-whitening  :                                           0.000000 |   HAC Trunct.Lag.        :     12.000000
================================================================================================================
```

(b) Kalman Smoother

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

In summary, the HP-Filter can reasonably accurately recover the transitory cyclical shock $\varepsilon_{2t} = \varepsilon_t^c$, but not the permanent shock. Moreover, the HP-Filter will produce negatively correlated trend-cycle shocks ($\text{Corr}(\varepsilon_{t|T}^*, \varepsilon_{t|T}^c) = -0.1907$), despite the data being generated from a model that has zero shock correlation. The HP-Filter shock correlation is thus spurious.