Documentation of R Code and Data for "Measuring the Natural Rate of Interest"*

Thomas Laubach John C. Williams

June 3, 2020

Note: This guide has been updated in May 2020 to include documentation for the COVID-adjusted LW model. Please see our note on adjusting the model to allow for continued estimation during and after the COVID-19 pandemic. Note that the code employs a toggle switch for 2020:Q1, and that standard error estimates will include ϕ beginning in 2020:Q2 (see final paragraph of linked note for more details).[1]

This note documents the R code used for the estimation of the natural rate of interest, potential GDP, and its trend growth rate for the United States presented in "Measuring the Natural Rate of Interest" (Laubach and Williams 2003; henceforth LW). It also catalogues steps to download and prepare data used in the estimation. The code documented here includes modifications described in the linked note to adjust for the COVID-19 pandemic, and uses an estimation sample of 1961:Q1 through 2020:Q1. **As of 2020:Q1, we also use a COVID-19 index. We will provide this index each quarter with the pre-processed data accompanying our updated estimates.** This index is derived from the Oxford COVID-19 Government Response Tracker's Stringency Index; see our note on COVID-adjusted models.

1 Code Layout and Directory Structure

There is one main R file, *run.lw.R*, which does the following:

1. Reads in pre-processed data to be used in the LW estimation;
2. Defines the sample period, constraints, and variables to be used throughout the estimation;

*Laubach: Board of Governors of the Federal Reserve System, Thomas.Laubach@frb.gov.

Williams: Federal Reserve Bank of San Francisco, John.C.Williams@sf.frb.org.

[1]"Adapting the Laubach and Williams and Holston, Laubach, and Williams Models to the COVID-19 Pandemic": https://www.newyorkfed.org/medialibrary/media/research/policy/rstar/LW_HLW_COVID_note

3. Runs the three-stage LW estimation;
4. Saves output.

This file calls multiple R functions and files, each of which are described in this guide. For reference, we are using R Version 3.3.1 at the time of release.

2 Basic Functions used Throughout LW Programs

In the accompanying set of code, these functions are stored in *utilities.R*.

Function: *shiftQuarter*
Description: This function takes in a (year, quarter) date in time series format and a shift number, and returns the (year, quarter) date corresponding to the shift. Positive values of shift produce leads and negative values of shift produce lags. For example, entering 2014q1 with a shift of -1 would return 2013q4. Entering 2014q1 with a shift of 1 would return 2014q2. In each case, the first argument of the function must be entered as a two-element vector, where the first element corresponds to the year and the second element corresponds to the quarter. For example, 2014q1 must be entered as "c(2014, 1)".

Function: *shiftMonth*
Description: This function takes in a (year, month) date in time series format and a shift number, and returns the (year, month) date corresponding to the shift. Positive values of shift produce leads and negative values of shift produce lags. For example, entering 2014m1 with a shift of -1 would return 2013m12. Entering 2014m1 with a shift of 1 would return 2014m2. In each case, the first argument of the function must be entered as a two-element vector, where the first element corresponds to the year and the second element corresponds to the month. This function is analogous to shiftQuarter().

Function: *gradient*
Description: This function computes the gradient of a function f given a vector input x.

3 R Packages

The "tis" package is used to manage time series data. We use the "nloptr" package for optimization.

4 Estimation

The results reported in Laubach and Williams (2003) are based on our estimation method described in Section 2 of the paper. The estimation proceeds in sequential steps through three stages, each of which is implemented in an R program. These and all other R programs are described in this section.

4.1 Main Estimation in *run.lw.R*

The program *run.lw.R* defines and trims the sample for the key variables: log output, inflation, the real and nominal short-term interest rates, and oil and import price inflation. It is also where the specified constraints on a_r and b_y are set. It calls the programs *rstar.stageX.R* to run the three stages of the LW estimation. Additionally, it calls the programs *median.unbiased.estimator.stageX.R* to obtain the signal-to-noise ratios λ_g and λ_z.

The programs *unpack.parameters.stageX.R* set up coefficient matrices for the corresponding state-space models for the given parameter vectors. In stages 2 and 3, we impose the constraints $a_r \leq -0.0025$ and $b_y \geq 0.025$. These constraints are labeled as a.r.constraint and b.y.constraint, respectively, in the code.

4.2 The Stage 1, 2, and 3 COVID-adjusted State-Space Models

This section presents the COVID-adjusted state-space models. See the linked note for modifications to the LW model during the COVID-19 pandemic. For reference, the next section presents the standard state-space models presented in our paper. The following section documents the corresponding R programs. Notation matches that of Hamilton (1994) and is also used in the R programs. All of the state-space models can be cast in the form:

$$\mathbf{y}_t = \mathbf{A}' \cdot \mathbf{x}_t + \mathbf{H}' \cdot \xi_t + \mathbf{v}_t \quad (1)$$

$$\xi_t = \mathbf{F} \cdot \xi_{t-1} + \mathbf{c} + \epsilon_t \quad (2)$$

Here, $\mathbf{y}_t$ is a vector of contemporaneous endogenous variables, while $\mathbf{x}_t$ is a vector of exogenous and lagged exogenous variables. ξ_t is vector of unobserved states. The vectors of stochastic disturbances $\mathbf{v}_t$ and ϵ_t are assumed to be Gaussian and mututally uncorrelated, with mean zero and covariance matrices $\mathbf{R}$ and $\mathbf{Q}$, respectively. The covariance matrix $\mathbf{R}$ is always assumed to be diagonal. $\mathbf{c}$ is 0 in Stages 2 and 3.

For each model, there is a corresponding vector of parameters to be estimated by maximum likelihood. Because maximum likelihood estimates of the innovations to g and z, σ_g and σ_z, are likely to be biased towards zero (see Section 2 of LW for explanation), we use Stock and Watson's (1998) mediun unbiased estimator to obtain estimates of two ratios, $\lambda_g \equiv \frac{\sigma_g}{\sigma_{y^*}}$ and $\lambda_z \equiv \frac{a_r \sigma_z}{\sigma_{\tilde{y}}}$. We impose these ratios when estimating the remaining model parameters by maximum likelihood.

4.2.1 The COVID-19 Model Adjustments

The direct effects of COVID on the economy are incorporated in the models as an adjustment to the natural rate of output. In particular, the adjusted natural rate of output is given by

$$y_t^* + \phi \frac{d_t}{100}$$

where ϕ is an estimated parameter that translates the COVID indicator into effects on output, and y_t^* is the standard natural rate of output. The output gap is correspondingly modified with the adjusted natural rate of output replacing the standard natural rate of output. That is, the modified output gap equals

$$100(y_t - y_t^*) - \phi d_t$$

where y_t is log real GDP. The adjusted output gap replaces the standard output gap in the measurement equations (the IS and Phillips curve equations). No other changes are made to the model specification.

4.2.2 The COVID-adjusted Stage 1 Model

The first-stage model, which corresponds to the *rstar.stage1.R* program, can be represented by the following matrices:

$$\mathbf{y}_t = [y_t, \pi_t]' \tag{3}$$

$$\mathbf{x}_t = \left[y_{t-1}, y_{t-2}, \pi_{t-1}, \pi_{t-2,4}, \pi_{t-5,8}, \pi_{t-1}^0 - \pi_{t-1}, \pi_t^m - \pi_t, d_t, d_{t-1}, d_{t-2}\right]' \tag{4}$$

$$\xi_t = \left[y_t^*, y_{t-1}^*, y_{t-2}^*\right]' \tag{5}$$

$$\mathbf{H}' = \begin{bmatrix} 1 & -a_1 & -a_2 \\ 0 & -b_3 & 0 \end{bmatrix}$$

$$\mathbf{A}' = \begin{bmatrix} a_1 & a_2 & 0 & 0 & 0 & 0 & 0 & \phi & -\phi a_1 & -\phi a_2 \\ b_3 & 0 & b_1 & b_2 & 1 - b_1 - b_2 & b_4 & b_5 & 0 & -\phi b_3 & 0 \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \sigma_{y^*}^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} g \\ 0 \\ 0 \end{bmatrix}$$

The vector of parameters to be estimated by maximum likelihood is as follows:

$$\theta_1 = [a_1, a_2, b_1, b_2, b_3, b_4, b_5, g, \sigma_1, \sigma_2, \sigma_4, \phi]$$

4.2.3 The COVID-adjusted Stage 2 Model

The second-stage model, which corresponds to the *rstar.stage2.R* program, can be represented by the following matrices:

$$\mathbf{y}_t = [y_t, \pi_t]' \tag{6}$$

$$\mathbf{x}_t = \left[y_{t-1}, y_{t-2}, r_{t-1}, r_{t-2}, \pi_{t-1}, \pi_{t-2,4}, \pi_{t-5,8}, \pi^0_{t-1} - \pi_{t-1}, \pi^m_t - \pi_t, 1, d_t, d_{t-1}, d_{t-2}\right]' \tag{7}$$

$$\xi_t = \left[y^*_t, y^*_{t-1}, y^*_{t-2}, g_{t-1}\right]' \tag{8}$$

$$\mathbf{H}' = \begin{bmatrix} 1 & -a_1 & -a_2 & a_5 \\ 0 & -b_3 & 0 & 0 \end{bmatrix}$$

$$\mathbf{A}' = \begin{bmatrix} a_1 & a_2 & \frac{a_3}{2} & \frac{a_3}{2} & 0 & 0 & 0 & 0 & 0 & a_4 & \phi & -\phi a_1 & -\phi a_2 \\ b_3 & 0 & 0 & 0 & b_1 & b_2 & 1-b_1-b_2 & b_4 & b_5 & 0 & 0 & -\phi b_3 & 0 \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \sigma^2_{y^*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\lambda_g \sigma_{y^*})^2 \end{bmatrix}$$

The vector of parameters to be estimated by maximum likelihood is as follows:

$$\theta_2 = [a_1, a_2, a_3, a_4, a_5, b_1, b_2, b_3, b_4, b_5, \sigma_1, \sigma_2, \sigma_4, \phi]$$

4.2.4 The COVID-adjusted Stage 3 Model

The third-stage model, which corresponds to the *rstar.stage3.R* program, can be represented by the following matrices:

$$\mathbf{y}_t = [y_t, \pi_t]' \tag{9}$$

$$\mathbf{x}_t = \left[y_{t-1}, y_{t-2}, r_{t-1}, r_{t-2}, \pi_{t-1}, \pi_{t-2,4}, \pi_{t-5,8}, \pi^0_{t-1} - \pi_{t-1}, \pi^m_t - \pi_t, d_t, d_{t-1}, d_{t-2}\right]' \tag{10}$$

$$\xi_t = \left[y^*_t, y^*_{t-1}, y^*_{t-2}, g_{t-1}, g_{t-2}, z_{t-1}, z_{t-2}\right]' \tag{11}$$

$$\mathbf{H}' = \begin{bmatrix} 1 & -a_1 & -a_2 & -c\frac{a_3}{2} & -c\frac{a_3}{2} & \frac{-a_3}{2} & \frac{-a_3}{2} \\ 0 & -b_3 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{A}' = \begin{bmatrix} a_1 & a_2 & \frac{a_3}{2} & \frac{a_3}{2} & 0 & 0 & 0 & 0 & 0 & \phi & -\phi a_1 & -\phi a_2 \\ b_3 & 0 & 0 & 0 & b_1 & b_2 & 1-b_1-b_2 & b_4 & b_5 & 0 & -\phi b_3 & 0 \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \left(1+\lambda_g^2\right)\sigma_{y^*}^2 & 0 & 0 & (\lambda_g \sigma_{y^*})^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (\lambda_g \sigma_{y^*})^2 & 0 & 0 & (\lambda_g \sigma_{y^*})^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \left(\frac{\lambda_z \sigma_{\tilde{y}}}{a_r}\right)^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The vector of parameters to be estimated by maximum likelihood is as follows:

$$\theta_3 = [a_1, a_2, a_3, b_1, b_2, b_3, b_4, b_5, c, \sigma_1, \sigma_2, \sigma_4, \phi]$$

4.3 The Standard Stage 1, 2, and 3 State-Space Models

This section includes the standard LW model presented in our paper, without the COVID-19 adjustment, for reference.

4.3.1 The Stage 1 Model

The first-stage model, which corresponds to the *rstar.stage1.R* program, can be represented by the following matrices:

$$\mathbf{y}_t = [y_t, \pi_t]' \tag{12}$$

$$\mathbf{x}_t = \left[y_{t-1}, y_{t-2}, \pi_{t-1}, \pi_{t-2,4}, \pi_{t-5,8}, \pi_{t-1}^0 - \pi_{t-1}, \pi_t^m - \pi_t\right]' \tag{13}$$

$$\xi_t = \left[y_t^*, y_{t-1}^*, y_{t-2}^*\right]' \tag{14}$$

$$\mathbf{H}' = \begin{bmatrix} 1 & -a_1 & -a_2 \\ 0 & -b_3 & 0 \end{bmatrix}, \quad \mathbf{A}' = \begin{bmatrix} a_1 & a_2 & 0 & 0 & 0 & 0 & 0 \\ b_3 & 0 & b_1 & b_2 & 1-b_1-b_2 & b_4 & b_5 \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \sigma_{y^*}^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} g \\ 0 \\ 0 \end{bmatrix}$$

The vector of parameters to be estimated by maximum likelihood is as follows:

$$\theta_1 = [a_1, a_2, b_1, b_2, b_3, b_4, b_5, g, \sigma_1, \sigma_2, \sigma_4]$$

4.3.2 The Stage 2 Model

The second-stage model, which corresponds to the *rstar.stage2.R* program, can be represented by the following matrices:

$$\mathbf{y}_t = [y_t, \pi_t]' \tag{15}$$

$$\mathbf{x}_t = \left[y_{t-1}, y_{t-2}, r_{t-1}, r_{t-2}, \pi_{t-1}, \pi_{t-2,4}, \pi_{t-5,8}, \pi^0_{t-1} - \pi_{t-1}, \pi^m_t - \pi_t, 1\right]' \tag{16}$$

$$\xi_t = \left[y^*_t, y^*_{t-1}, y^*_{t-2}, g_{t-1}\right]' \tag{17}$$

$$\mathbf{H}' = \begin{bmatrix} 1 & -a_1 & -a_2 & a_5 \\ 0 & -b_3 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}' = \begin{bmatrix} a_1 & a_2 & \frac{a_3}{2} & \frac{a_3}{2} & 0 & 0 & 0 & 0 & 0 & a_4 \\ b_3 & 0 & 0 & 0 & b_1 & b_2 & 1 - b_1 - b_2 & b_4 & b_5 & 0 \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \sigma^2_{y^*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\lambda_g \sigma_{y^*})^2 \end{bmatrix}$$

The vector of parameters to be estimated by maximum likelihood is as follows:

$$\theta_2 = [a_1, a_2, a_3, a_4, a_5, b_1, b_2, b_3, b_4, b_5, \sigma_1, \sigma_2, \sigma_4]$$

4.3.3 The Stage 3 Model

The third-stage model, which corresponds to the *rstar.stage3.R* program, can be represented by the following matrices:

$$\mathbf{y}_t = [y_t, \pi_t]' \tag{18}$$

$$\mathbf{x}_t = \left[y_{t-1}, y_{t-2}, r_{t-1}, r_{t-2}, \pi_{t-1}, \pi_{t-2,4}, \pi_{t-5,8}, \pi^0_{t-1} - \pi_{t-1}, \pi^m_t - \pi_t\right]' \tag{19}$$

$$\xi_t = \left[y^*_t, y^*_{t-1}, y^*_{t-2}, g_{t-1}, g_{t-2}, z_{t-1}, z_{t-2}\right]' \tag{20}$$

$$\mathbf{H}' = \begin{bmatrix} 1 & -a_1 & -a_2 & -c\frac{a_3}{2} & -c\frac{a_3}{2} & \frac{-a_3}{2} & \frac{-a_3}{2} \\ 0 & -b_3 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}' = \begin{bmatrix} a_1 & a_2 & \frac{a_3}{2} & \frac{a_3}{2} & 0 & 0 & 0 & 0 & 0 \\ b_3 & 0 & 0 & 0 & b_1 & b_2 & 1-b_1-b_2 & b_4 & b_5 \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \left(1+\lambda_g^2\right)\sigma_{y^*}^2 & 0 & 0 & (\lambda_g\sigma_{y^*})^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (\lambda_g\sigma_{y^*})^2 & 0 & 0 & (\lambda_g\sigma_{y^*})^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \left(\frac{\lambda_z\sigma_{\tilde{y}}}{a_r}\right)^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The vector of parameters to be estimated by maximum likelihood is as follows:

$$\theta_3 = [a_1, a_2, a_3, b_1, b_2, b_3, b_4, b_5, c, \sigma_1, \sigma_2, \sigma_4]$$

4.4 R Programs to Run the State-Space Models

The programs *rstar.stageX.R* run the models in stages 1-3 of the LW estimation.

4.5 R Programs for Median Unbiased Estimators

The function *median.unbiased.estimator.stage1.R* computes the exponential Wald statistic of Andrews and Ploberger (1994) for a structural break with unknown break date from the first difference of the preliminary estimate of the natural rate of output from the stage 1 model to obtain the median unbiased estimate of λ_g.

The function *median.unbiased.estimator.stage2.R* applies the exponential Wald test for an intercept shift in the IS equation at an unknown date to obtain the median unbiased estimate of λ_z, taking as input estimates from the stage 2 model.

4.6 Kalman Filter Programs

Within the program *kalman.states.R*, the function kalman.states() calls kalman.states.filtered() and kalman.states.smoothed() to apply the Kalman filter and smoother. It takes as input the coefficient matrices for the given state-space model as well as the conditional expectation and covariance matrix of the initial state, xi.tm1tm1 $\left(\xi_{t-1|t-1}\right)$ and P.tm1tm1 $\left(P_{t-1|t-1}\right)$, respectively. *kalman.states.wrapper.R* is a wrapper function for *kalman.states.R* that specifies inputs based on the estimation stage.

4.7 Log Likelihood Programs

The function *kalman.log.likelihood.R* takes as input the coefficient matrices of the given state-space model and the conditional expectation and covariance matrix of the initial state and returns the log likelihood value and a vector with the log likelihood at each time t. *log.likelihood.wrapper.R* is a wrapper function for *kalman.log.likelihood.R* that specifies inputs based on the estimation stage.

4.8 Standard Error Program

The function *kalman.standard.errors.R* computes confidence intervals and corresponding standard errors for the estimates of the states using Hamilton's (1986) Monte Carlo procedure that accounts for both filter and parameter uncertainty. See footnote 7 in LW.

4.9 Miscellaneous Programs

The function *calculate.covariance.R* calculates the covariance matrix of the initial state from the gradients of the likelihood function. The function *format.output.R* generates a dataframe to be written to a CSV containing one- and two-sided estimates, parameter values, standard errors, and other statistics of interest.

5 References

Andrews, Donald, and Werner Ploberger. 1994. "Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative," *Econometrica*, 62, 1383-1414.

Fagan, Gabriel, Jerome Henry, and Ricardo Mestre. 2001. "An Area-Wide Model (AWM) for the Euro Area," European Central Bank Working Paper Series No. 42, January.

Hamilton, James D. 1986. "A Standard Error for the Estimated State Vector of a State-Space Model," *Journal of Econometrics,* 33, 387-397.

Hamilton, James D. 1994. *Time Series Analysis,* Princeton, NJ: Princeton University Press.

Laubach, Thomas, and John C. Williams. 2003. "Measuring the Natural Rate of Interest," *Review of Economics and Statistics,* 85(4), November, 1063-1070.

Stock, James, and Mark Watson. 1998. "Median Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model," *Journal of the American Statistical Association*, 93, March, 349–358.