Appendix to

"Using Advance Layoff Notices as a Labor Market Indicator"

by Pawel M. Krolikowski, Kurt G. Lunsford, and Meifeng Yang

Federal Reserve Bank of Cleveland, *Economic Commentary*, 2019-21

This appendix provides some details of the dynamic factor model (DFM) and the regressions used to produce Table 1.

The DFM is as follows. Let $WARN_{s,t}$ denote the number of workers affected by WARN notices in state s and in month t. As noted in the *Commentary*, these data are seasonally adjusted. Then, define $z_{s,t} = \ln(WARN_{s,t})$ and $z_t = [z_{1,t}, \dots, z_{N,t}]'$, in which N is the number of states in the data. We treat $z_{s,t}$ as unobserved if $WARN_{s,t} = 0$. The DFM takes the following structure:

$$z_t = d + \Lambda_t f_t + \epsilon_t,$$

in which f_t is the scalar national WARN factor. In addition, $d = [d_1, \dots, d_N]'$ is an $N \times 1$ vector, $\Lambda_t = [\lambda_{1,t}, \dots, \lambda_{N,t}]'$ is an N-dimensional process of factor loadings, and $\epsilon_t = [\epsilon_{1,t}, \dots, \epsilon_{N,t}]'$ is an N-dimensional process of state-specific shocks. We assume that ϵ_t is an independent and identically distributed multivariate normal process with mean zero and a diagonal covariance matrix.

Because labor market data often display some persistence, we assume that the WARN factor follows an AR(1) process:

$$f_t = Af_{t-1} + \eta_t,$$

with $|A| < 1$ and in which η_t is an independent and identically distributed normal process with mean zero. We allow for $A = 0$, implying that the AR(1) assumption does not impose persistence on f_t.

The above equations compose our DFM. We estimate d_s by taking the average of $z_{s,t}$ over the sample in which $z_{s,t}$ is observed. We then subtract d from z_t and use an expectation maximization (EM) algorithm to estimate the covariance matrix of ϵ_t, the value of A, and the variance of η_t by maximum likelihood. We then use these maximum likelihood estimates to produce an estimate of the WARN factor, $\{\hat{f}_t\}_{t=1}^{T}$. We follow the EM algorithms in Shumway and Stoffer (1982) and Bańbura and Modugno (2014), which allow for some observations of $z_{s,t}$ to be unobserved. We produce estimates of the WARN factor from a Kalman filter and smoother, implying that current information is incorporated into the estimates of past values of the WARN factor. That is, $z_{s,T}$ is allowed to affect the estimate of f_τ for $\tau \leq T$. As part of the EM algorithm, we impose that the unconditional variance of f_t equals 1. We do not estimate Λ_t. Rather, let $E_{s,t}$ denote the level of employment in state s and in month t. Then, we impose that Λ_t is proportional to $\left[\ln\left(E_{1,t-1}\right), \ldots, \ln\left(E_{N,t-1}\right)\right]' / \sum_{s=1}^{50} \ln\left(E_{s,t-1}\right)$.[1] We impose these loadings so that the DFM puts more weight on larger states when estimating f_t with the intent of having the WARN factor be nationally representative. Following the spirit of Solon, Haider, and Wooldridge (2015), our intent is to make the WARN factor "representative of the target population," which is the whole United States.

Before discussing the regressions in Table 1, we produce mean squared errors for the estimates of f_t. We do this using a parametric bootstrap. Within each bootstrap loop, we simulate $\{\epsilon_t\}_{t=1}^{T}$ and $\{\eta_t\}_{t=1}^{T}$ from the distributions implied by the maximum likelihood estimates. Then, we simulate a value of f_0 from its unconditional distribution and use these simulated variables along with the estimate of d and Λ_t to create simulated values of $\{f_t\}_{t=1}^{T}$ and $\{z_t\}_{t=1}^{T}$. Then, we re-estimate d and re-run the EM algorithm with the simulated values of $\{z_t\}_{t=1}^{T}$. In this process, we impose the same pattern of missing data on the simulated values of $\{z_t\}_{t=1}^{T}$ that exists for the actual values of $\{z_t\}_{t=1}^{T}$. The last step in each bootstrap loop is to use the bootstrapped maximum likelihood parameters from the EM algorithm to produce a bootstrapped $\{f_t\}_{t=1}^{T}$. We run 500 bootstrap replications and compute the mean squared errors of $\{f_t\}_{t=1}^{T}$ using equations (7) and (8) in Pfeffermann and Tiller (2005). We denote these mean squared errors with $\sigma_t^2 = E\left(\hat{f}_t - f_t\right)^2$, and our parametric bootstrap gives us an estimate of $\{\sigma_t^2\}_{t=1}^{T}$.

For the regressions in Table 1, UI_t denotes national initial unemployment insurance (UI) claims in month t,[2] $\Delta U_t = U_t - U_{t-1}$ denotes the change in the national unemployment rate in month t, and $\Delta E_t = E_t - E_{t-1}$ denotes the change in the level of national private employment in month t. For notation, define $X_t = [UI_t, \Delta U_t, \Delta E_t]'$. We estimate three regressions for Table 1:

$$UI_t = \beta_0 + \sum_{i=1}^{p} X_{t-i}'\beta_i + \sum_{j=1}^{q} \hat{f}_{t-j}\gamma_j + v_t,$$

$$\frac{(\Delta U_{t+2} + \Delta U_{t+1} + \Delta U_t)}{3} = \beta_0 + \sum_{i=1}^{p} X_{t-i}'\beta_i + \sum_{j=1}^{q} \hat{f}_{t-j}\gamma_j + v_{t+2},$$

and

$$\frac{(\Delta E_{t+2} + \Delta E_{t+1} + \Delta E_t)}{3} = \beta_0 + \sum_{i=1}^{p} X_{t-i}'\beta_i + \sum_{j=1}^{q} \hat{f}_{t-j}\gamma_j + v_{t+2}.$$

[1] We already impose that f_t has an unconditional variance of 1. This normalization implies that we can impose Λ_t only up to proportion, but we have to let the scale of Λ_t adjust to allow for f_t having an unconditional variance of 1.

[2] We use monthly averages of the weekly UI data from the FRED database in our regressions.

In all equations, we use $p = 3$ and $q = 3$. We estimate one equation at a time so that the βs and γs can be different for each equation. We use ordinary least squares and treat $\{\hat{f}_t\}_{t=1}^{T}$ as the true value of $\{f_t\}_{t=1}^{T}$ when estimating these coefficients. We show the estimates of the γs in Table 1.

We adjust the standard errors and p-values of the Wald tests to account for the fact that we use $\{\hat{f}_t\}_{t=1}^{T}$ and not $\{f_t\}_{t=1}^{T}$. We do this following the spirit of Murphy and Topel (1985). Consider the general linear regression

$$y_t = X_t'\beta + F_t'\gamma + v_t,$$

in which X_t is a K-dimensional process of control variables and $F_t = [f_{t-1}, \dots, f_{t-q}]'$. Then, the estimates of the regression coefficients are given by

$$\begin{bmatrix}\hat{\beta} \\ \hat{\gamma}\end{bmatrix} = \left(\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} [X_t' \quad \hat{F}_t']\right)^{-1} \left(\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} y_t\right),$$

in which $\hat{F}_t = [\hat{f}_{t-1}, \dots, \hat{f}_{t-q}]'$. The above two equations imply

$$T^{\frac{1}{2}}\left(\begin{bmatrix}\hat{\beta} \\ \hat{\gamma}\end{bmatrix} - \begin{bmatrix}\beta \\ \gamma\end{bmatrix}\right) = \left(T^{-1}\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} [X_t' \quad \hat{F}_t']\right)^{-1} \left(T^{-\frac{1}{2}}\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} \left((F_t' - \hat{F}_t')\gamma + v_t\right)\right).$$

We assume that $T^{-1}\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} [X_t' \quad \hat{F}_t']$ converges in probability to a matrix Q and that $T^{-\frac{1}{2}}\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} \left((F_t' - \hat{F}_t')\gamma + v_t\right)$ converges in distribution to a multivariate normal with mean zero and covariance matrix Ω.

We now discuss the estimation of Ω. We assume that $\Omega = \Omega_f + \Omega_v$, in which $T^{-\frac{1}{2}}\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} (F_t' - \hat{F}_t')\gamma$ converges in distribution to a multivariate normal with mean zero and covariance Ω_f and $T^{-\frac{1}{2}}\sum_{t=q+1}^{T} \begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} v_t$ converges in distribution to a multivariate normal with mean zero and covariance Ω_v. We estimate Ω_v in the standard way using the Bartlett kernel with a truncation parameter of 4 (Newey and West, 1987). Ω_v is the conventional long-run variance of $\begin{bmatrix}X_t \\ \hat{F}_t\end{bmatrix} v_t$, and it would be sufficient for inference if $\hat{f}_t$ were not a generated regressor. However, because $\hat{f}_t$ is a generated regressor, we also compute Ω_f, which we now turn to.

For Ω_f, note that $F_t' - \hat{F}_t' = \left[\left(f_{t-1} - \hat{f}_{t-1}\right), \ldots, \left(f_{t-q} - \hat{f}_{t-q}\right)\right]$, $f_t - \hat{f}_t$ is serially uncorrelated, the expectation of $\left(f_t - \hat{f}_t\right)^2$ is equal to σ_t^2, which we estimate above with the parametric bootstrap, and $\begin{bmatrix} X_t \\ \hat{F}_t \end{bmatrix} (F_t' - \hat{F}_t')\gamma = \begin{bmatrix} X_t \\ \hat{F}_t \end{bmatrix} \left[\left(f_{t-1} - \hat{f}_{t-1}\right)\gamma_1 + \cdots + \left(f_{t-q} - \hat{f}_{t-q}\right)\gamma_q\right]$. Then, we compute

$$\hat{\Gamma}_0 = T^{-1} \sum_{t=q+1}^{T} \begin{bmatrix} X_t \\ \hat{F}_t \end{bmatrix} [X_t' \quad \hat{F}_t'] \left(\hat{\sigma}_{t-1}^2 \hat{\gamma}_1^2 + \cdots + \hat{\sigma}_{t-q}^2 \hat{\gamma}_q^2\right),$$

$$\hat{\Gamma}_1 = T^{-1} \sum_{t=q+2}^{T} \begin{bmatrix} X_t \\ \hat{F}_t \end{bmatrix} [X_{t-1}' \quad \hat{F}_{t-1}'] \left(\hat{\sigma}_{t-2}^2 \hat{\gamma}_1 \hat{\gamma}_2 + \cdots + \hat{\sigma}_{t-q}^2 \hat{\gamma}_{q-1} \hat{\gamma}_q\right),$$

$$\vdots$$

$$\hat{\Gamma}_{q-1} = T^{-1} \sum_{t=2q}^{T} \begin{bmatrix} X_t \\ \hat{F}_t \end{bmatrix} [X_{t-q+1}' \quad \hat{F}_{t-q+1}'] \hat{\sigma}_{t-q}^2 \hat{\gamma}_1 \hat{\gamma}_q,$$

in which $\hat{\sigma}_t^2$ is the bootstrapped estimate of σ_t^2 and $\hat{\gamma}_j$ is the ordinary least squares estimate of γ_j. Then, we compute $\hat{\Omega}_f = \hat{\Gamma}_0 + k(1/q)\left(\hat{\Gamma}_1 + \hat{\Gamma}_1'\right) + \cdots + k((q-1)/q)\left(\hat{\Gamma}_{q-1} + \hat{\Gamma}_{q-1}'\right)$, in which $k(j/q) = (q - j)/q$ is the Bartlett kernel.

For standard errors, we compute $\hat{V} = \hat{Q}^{-1}\hat{\Omega}\hat{Q}^{-1} = \hat{Q}^{-1}\left(\hat{\Omega}_f + \hat{\Omega}_v\right)\hat{Q}^{-1}$, divide by $T - 1 - 3p - q$, and take the square root of diagonal elements. Because $p = 3$ and $q = 3$, we divide by $T - 13$. We do this to match the degrees-of-freedom adjustment in Stata. For Wald statistics in Table 1, consider the selection matrix R such that $\hat{\gamma} = R \begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix}$. Then, define $\hat{V}_\gamma = R\hat{V}R'$. The Wald statistic is $(T - 1 - 3p - q)\hat{\gamma}'\hat{V}_\gamma^{-1}\hat{\gamma}$, which we treat as distributed χ_q^2 to compute the p-values.

References

Bańbura, Marta and Michele Modugno (2014). "Maximum Likelihood of Factor Models on Datasets with Arbitrary Pattern of Missing Data." *Journal of Applied Econometrics,* November 2014, 29(1): 133-160.

Murphy, Kevin M. and Robert H. Topel (1985). "Estimation and Inference in Two-Step Econometric Models." *Journal of Business & Economic Statistics,* October 1985, 3(4): 88-97.

Newey, Whitney K. and Kenneth D. West (1987). "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica,* May 1987, 55(3): 703-708.

Pfeffermann, Danny and Richard Tiller (2005). “Bootstrap Approximation to Prediction MSE for State-Space Models with Estimated Parameters.” *Journal of Time Series Analysis*, November 2006, 26(6): 893-916.

Shumway, Robert H., and David S. Stoffer (1982). “An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm.” *Journal of Times Series Analysis*, July 1982, 3(4): 253-264.

Solon, Gary, and Steven J. Haider, and Jeffrey M. Wooldridge (2015). “What Are We Weighting For?” *Journal of Human Resources*, March 2015, 50(2): 301-316.