LP-VAR Simulation: Technical Note

1 Overview of simulation experiments

We consider a large-scale Dynamic Factor Model (DFM) as the encompassing model. We then randomly draw five variables from the encompassing model as one DGP to simulate data and repeat it thousands of times. For each DGP, we consider three estimands: IRF identified by observed shocks, IRF identified by IV, IRF identified by recusive shocks. For each estimand, we compare the performance of various different estimation methods, including LP, VAR and their variants.

1.1 The Encompassing Model

We use the dynamic factor model (DFM) of Stock and Watson (2016). That model takes the form

$$
\begin{aligned}
f_t &= \Phi_1 f_{t-1} + \Phi_2 f_{t-2} + H\epsilon_t && (1.1)\\
X_t &= \Lambda f_t + v_t && (1.2)\\
v_{it} &= \Delta_{i1} v_{i,t-1} + \Delta_{i2} v_{i,t-2} + \Xi_i \xi_{it} && (1.3)
\end{aligned}
$$

where Δ_1, Δ_2, Ξ are all diagonal.

1.2 IV DGP

IV has the followin DGP:

$$z_t = \rho_z z_{t-1} + \alpha \mathbf{w}' \epsilon_t + \nu_t \tag{1.4}$$

where $\nu_t \sim \mathcal{N}(0, \sigma_\nu^2)$, where $\mathbf{w}$ is a fixed shock weights vector.

2 Transforming Representation

2.1 General ABCDEF Representation

The ABCDEF representation takes the form of:

$$s_t = A s_{t-1} + B\epsilon_t \tag{2.1}$$

$$
\begin{aligned}
y_t &= Cs_{t-1} + D\epsilon_t + e_t^* && (2.2)\\
e_t &= Ee_{t-1} + F\omega_t && (2.3)
\end{aligned}
$$

where $(\epsilon_t, \omega_t) \sim \mathcal{N}(0, I)$, $e_t = (e_t^{*\prime}, e_{t-1}^{*}{}', \cdots, e_{t-\ell+1}^{*}{}')'$, and ℓ is lag order in measurement errors.

The mapping from DFM into the generic ABCDEF framework is straightforward:

$$
\begin{aligned}
\begin{pmatrix} f_t \\ f_{t-1} \end{pmatrix} &= \underbrace{\begin{pmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{pmatrix}}_{A} \begin{pmatrix} f_{t-1} \\ f_{t-2} \end{pmatrix} + \underbrace{\begin{pmatrix} H \\ 0 \end{pmatrix}}_{B} \epsilon_t && (2.4)\\
X_t &= \underbrace{\begin{pmatrix} \Lambda & 0 \end{pmatrix} \begin{pmatrix} \Phi_1 & \Phi_2 \\ I & 0 \end{pmatrix}}_{C} \begin{pmatrix} f_{t-1} \\ f_{t-2} \end{pmatrix} + \underbrace{\begin{pmatrix} \Lambda & 0 \end{pmatrix} \begin{pmatrix} H \\ 0 \end{pmatrix}}_{D} \epsilon_t + v_t && (2.5)\\
\begin{pmatrix} v_t \\ v_{t-1} \end{pmatrix} &= \underbrace{\begin{pmatrix} \Delta_1 & \Delta_2 \\ I & 0 \end{pmatrix}}_{E} \begin{pmatrix} v_{t-1} \\ v_{t-2} \end{pmatrix} + \underbrace{\begin{pmatrix} \Xi \\ 0 \end{pmatrix}}_{F} \xi_t && (2.6)
\end{aligned}
$$

2.2 Degree of invertibility

To compute the degree of invertibility, we first need to get an ABCD representation of the following form:

$$
\begin{aligned}
s_t &= As_{t-1} + B\epsilon_t && (2.7)\\
y_t &= Cs_t + D\xi_t && (2.8)
\end{aligned}
$$

In our DFM model, note that:

$$
\underbrace{(I - \Delta(L))y_t}_{y_t^*} = (I - \Delta(L))\Lambda(I - \Phi(L))^{-1}H\epsilon_t + \Xi\omega_t
$$

Now let $\zeta_t = (\epsilon_t', \epsilon_{t-1}', \cdots, \epsilon_{t-H}')'$, where H is sufficiently large. The mapping from DFM to this ABCD form is derived as follows:

$$
\begin{aligned}
\zeta_t &= \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix} \zeta_{t-1} + \begin{pmatrix} I \\ 0 \end{pmatrix} \epsilon_t\\
y_t^* &= C\zeta_t + \Xi\xi_t
\end{aligned}
$$

where $(I - \Delta(L))\Lambda(I - \Phi(L))^{-1}H\epsilon_t \approx C\zeta_t$.

We can then run a long Kalman filter algorithm until convergence to compute $\text{Var}(\zeta_t|y_t^*, \cdots y_{t-\infty}^*)$, which equals $\text{Var}(\zeta_t|y_t, \cdots y_{t-\infty})$. The degree of invertibility of the true shock is then easily recovered as $\text{Var}(\mathbf{w}'\epsilon_t|y_t, \cdots y_{t-\infty})$.

2.3 Recursive IRF estimand

To derive the true IRF to the recursive shock, we need to first get an ABCD representation consistent with Fernandez-Villaverde et al. (2005):

$$
\begin{aligned}
s_{t+1} &= As_t + Bw_t && (2.9)\\
y_t &= Cs_t + Dw_t && (2.10)
\end{aligned}
$$

We can rewrite the DFM in the form consistent with Fernandez-Villaverde et al. (2005) as follows:

$$
\underbrace{\begin{pmatrix} I & -\Delta_1 & -\Delta_2 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \end{pmatrix}}_{y_t^*} = \begin{pmatrix} I & -\Delta_1 & -\Delta_2 \end{pmatrix} \begin{pmatrix} \Lambda & & \\ & \Lambda & \\ & & \Lambda \end{pmatrix} \begin{pmatrix} f_t \\ f_{t-1} \\ f_{t-2} \end{pmatrix} + \begin{pmatrix} 0 & \Xi \end{pmatrix} \begin{pmatrix} \epsilon_{t+1} \\ \xi_t \end{pmatrix}
$$

$$
\underbrace{\begin{pmatrix} f_{t+1} \\ f_t \\ f_{t-1} \end{pmatrix}}_{s_{t+1}} = \begin{pmatrix} \Phi_1 & \Phi_2 & 0 \\ I & 0 & 0 \\ 0 & I & 0 \end{pmatrix} \underbrace{\begin{pmatrix} f_t \\ f_{t-1} \\ f_{t-2} \end{pmatrix}}_{s_t} + \begin{pmatrix} H & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} \epsilon_{t+1} \\ \xi_t \end{pmatrix}}_{w_t}
$$

Then we can use Fernandez-Villaverde et al. (2005) to transform this ABCD representation to a VAR(∞) representation as $y_t^* = \Psi^*(L)y_t^* + u_t$ (See details in Section A.4 in our paper). This in particular gives the representation and IRFs for $y_t^* = (1 - \Delta(L))y_t$. After that, we can easily transform its IRFs to get the IRFs of y_t.

2.4 Persistency of Observables

The first of our population persistency measures is defined as the ratio between the trace of long-run variance over the trace of simple variance, for each specification. Simple variance and long-run variance are calculated as follows:

- Simple Variance. Using the generic ABCDEF form in Section 2.1, we can first separately compute $Var(s_t)$, $Var(\epsilon_t)$, and $Var(e_t^*)$. Then we combine these results together to compute $Var(y_t)$.
- Long-run Variance. Similar to Section 2.3, we can derive the VAR(∞) representation as $y_t = \Psi(L)y_{t-1} + u_t$ and the Wold representation as $y_t = \theta(L)u_t$, where u_t is the reduced form

error. Then we can derive long-run variance $LRV(y_t) = \theta(1) \cdot LRV(u_t) \cdot \theta(1)'$.

The second persistency meansure is computing the largest root. Since we have got the VAR(∞) representation of y_t, we can truncate this VAR(∞) at a large lag order (Currently we truncate at 50 lags). After that, we compute the largest root in this truncated VAR.

Also, we have our population measure for the model fit of VAR(p), where we compute the relative fraction of the coefficients beyond lag p in the above-mentioned truncated VAR, i.e. $\sum_{l=p+1}^{50} ||A_l|| / \sum_{l=1}^{50} ||A_l||$, as a measure of VAR(p) fit. The matrix norm we use is Frobenius norm.

2.5 IV strength

To compute the population measure of the IV strength, we need to first incorporate the IV DGP into the DFM to get an augmented DFM as:

$$
\begin{aligned}
\underbrace{\begin{pmatrix} f_t \\ z_t \end{pmatrix}}_{f_t^*} &= \begin{pmatrix} \Phi_1 & 0 \\ 0 & \rho_z \end{pmatrix} \begin{pmatrix} f_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \Phi_2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} f_{t-2} \\ z_{t-2} \end{pmatrix} + \begin{pmatrix} H & 0 \\ \alpha \mathbf{w}' & \sigma_v \end{pmatrix} \underbrace{\begin{pmatrix} \epsilon_t \\ \nu_t \end{pmatrix}}_{\epsilon_t^*} \\
\underbrace{\begin{pmatrix} \bar{w}_t \\ z_t \end{pmatrix}}_{w_t} &= \begin{pmatrix} \Lambda & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} f_t \\ z_t \end{pmatrix} + \begin{pmatrix} v_t \\ \tilde{v}_t \end{pmatrix} \\
\begin{pmatrix} v_t \\ \tilde{v}_t \end{pmatrix} &= \begin{pmatrix} \Delta_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_{t-1} \\ \tilde{v}_{t-1} \end{pmatrix} + \begin{pmatrix} \Delta_2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_{t-2} \\ \tilde{v}_{t-2} \end{pmatrix} + \begin{pmatrix} \Xi & 0 \\ 0 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} \xi_t \\ \tilde{\xi}_t \end{pmatrix}}_{\xi_t^*}
\end{aligned}
$$

where $\tilde{v}_t$ and $\tilde{\xi}_t$ are just meaningless placeholders to make the model consistent with the standard DFM notation.

Similar to the steps in Section 2.3, we can transform the augmented DFM to the VAR(∞) representation as $w_t = \Psi^w(L) w_{t-1} + u_t^w$. We then straightforwardly get the residual forecasting variance of the normalization variable y_t after controlling lagged z_t and lagged $\bar{w}_t$, i.e. $\mathrm{Var}(y_t | w_{t-1}, \cdots, w_{t-\infty})$. After that, we compute how much variance in y_t can be explained by z_t after controlling for lagged w_t. To do so we define the residual $u_{i,t} \equiv y_t | (w_{t-1}, \cdots, w_{t-\infty})$. Note that, by Frisch-Waugh Theorem, the explained variance by z_t is

$$
\begin{aligned}
& \mathrm{Var}(y_t | w_{t-1}, \cdots, w_{t-\infty}) - \mathrm{Var}(y_t | z_t, w_{t-1}, \cdots, w_{t-\infty}) \\
= & \mathrm{Var}(u_{i,t}) - \mathrm{Var}(u_{i,t} | \alpha \mathbf{w}' \epsilon_t + \sigma_v \nu_t) \\
= & \mathrm{Var}[\mathrm{E}(u_t | \alpha \mathbf{w}' \epsilon_t + \sigma_v \nu_t)]
\end{aligned}
$$

$$= \quad \mathrm{IRF}_{y_i,\mathbf{w}'\varepsilon_t}(0)^2 \frac{\alpha^2}{\alpha^2+\sigma_v^2}$$

where $\mathrm{IRF}_{y,\mathbf{w}'\varepsilon_t}(0)$ is the impact response of y_t to the true shock. Finally, the population IV-strength is just the ratio between $\mathrm{Var}(y_t|w_{t-1},\cdots,w_{t-\infty})-\mathrm{Var}(y_t|z_t,w_{t-1},\cdots,w_{t-\infty})$ and $\mathrm{Var}(y_t|w_{t-1},\cdots,w_{t-\infty})$, i.e. the fraction of residual variance explained by IV.